The genotypic and phenotypic features of familial tooth agenesis in consanguineous families

SHUROG ABDULRAHMAN ALFAWAZ

BDS, MClinDent (Paed Dent, University of London), MPaedDent RCS(Eng)

Supervisors: Professor Ferranti Wong
Professor David Kelsell

Thesis submitted in fulfilment of the requirements for the degree of Doctor of Philosophy in the Faculty of Medicine, University of London

Centre for Oral Growth and Development
Institute of Dentistry
Barts and the London School of Medicine and Dentistry,
Queen Mary, University of London

April 2013

I dedicate this research work with love to

my parents,

my husband

Dr. Mesheal Alrobian

and my wonderful daughters

Lulu and Nora

Barts and The London
School of Medicine and Dentistry

Declaration regarding plagiarism

I declare that the coursework material attached herewith is entirely my own work and that I have attributed any brief quotations both at the appropriate point in the text and in the bibliography at the end of this piece of work.

I also declare that I have not used extensive quotations or close paraphrasing and that I have neither copied from the work of another person, nor used the ideas of another person, without proper acknowledgement.

Name: Shurog Abdulrahman Alfawaz Course: PhD

Title of work submitted:
The Genotypic and Phenotypic Features of Familial Tooth Agenesis in Consanguineous Families

Examination: A thesis submitted for the degree of Doctor of Philosophy, University of London

Signature:  Date:25/04/2012

Abstract

An evolution in methods of identifying the causal mutations and candidate genes for Mendelian disorders has occurred recently. Though several studies have reported the causative mutations in non syndromic tooth agenesis, there are only two reports on autosomal recessive nonsyndromic tooth agenesis in consanguineous families. Since the consanguinity rate is high in tribal population of Saudi and Pakistan, this study identified consanguineous families in these populations to investigate the molecular basis of tooth agenesis.

This study aimed to study the phenotype of familial tooth agenesis in Saudi and Pakistani families of consanguineous marriage, and to identify the causative mutations. A further aim was to investigate the influence of the agenesis gene on tooth size in one large Saudi family.

Sixteen families with non-syndromic tooth agenesis, were clinically characterised for the pattern of inheritance, tooth agenesis severity and type of teeth affected. Genetic analysis including homozygosity mapping and exome sequencing was performed in eight families. Hand measurements of Saudi family tooth dimensions were made on dental study models.

The most affected tooth type in studied families were the lower second premolar and upper lateral incisors. The homozygosity mapping approach failed to identify the regions of the diseased mutations in these families. The exome sequencing data revealed a heterozygous novel frameshift mutation in exon 2 of the MSX1 gene, c.750_751insACCGGCTGCC (p.A250fs) in one Saudi family and a homozygous novel mutation in exon 8 of the SMOC2

gene, c.681T>A (p.C227X), in the Pakistani family. It was found that there was a significant ($p < .0.05$) trend of reduction in the crown sizes both in the tooth agenesis group and their family members with no tooth agenesis compared to the control group.

It is concluded that homozygosity mapping was not sensitive to identify the elusive tooth agenesis gene and whole exome sequencing technique is needed in future studies. The tooth measurement study indicates and confirms that crown size is closely related to tooth agensis.

Acknowledgements

I would like to thank my academic supervisors Professor Ferranti Wong and Professor David Kelsell for their guidance and continued support during my PhD project. I would especially like to express my sincerest gratitude to Professor David Kelsell, without his guidance and commitment this project would not have been possible. He provided me with unwavering support and encouragement for which I am truly grateful.

I am deeply thankful to my previous supervisor, Professor Mark Hector, who is now the Dean of Dentistry at the University of Dundee for his great help and support.

I also would like to thank Professor Kelsell's group at the Centre for Cutaneous Research, Blizard Institute of Cell and Molecular Science. A special thanks to Dr Diana Blaydon and Dr Claire Scott for all their laboratory and technical training and to Dr Vincent Plagnol, Dr Charles Mein and Dr Monika Bozek for their valuable contribution to this project.

I wish to thank Professor Alan Hennery Brook who was generous with his time and advice regarding the initial tooth dimensions study, to Dr Janice Fearne and Ms. Francesca Fong for their help and cooperation for identifying the Pakistan family with the *SMOC2* genetic defect.

I am deeply thankful to Dr Majed Aljuraisy, the head of the Research Office at King Abdullah International Medical Research Centre, to Professor Mostafa Aboulforoh and to Dr Reem Albanyan for their great help and cooperation during the application for ethical approval of my research in Saudi Arabia. I also wish to thank Dr Firas Alkhayal and Dr Rahsa Al Sanie for their help in sample collection at King Abdul-Aziz Medical City in Riyadh, Saudi Arabia.

I express my special thanks to the members of the families who participated in this project.

I am deeply thankful to my great friends Dr Hanadi Lingawi, Miss Areej Asiery and Dr Amal Sindi for their valuable support and advice throughout my postgraduate study.

My deep appreciation and gratitude go to the Saudi Ministry of Higher Education and the Saudi Cultural Bureau in the UK for their financial grant and their continuous support throughout my course of studies.

Thanks to my family for their sustained love and support.

Table of Contents

List of Figures

List of Tables

List of abbreviations

µl	Microlitre
°C	Degree Celsius
3D	Three dimensional
A	Adenine
AD	Alzheimer disease
AEC	Ankyloblepharon-ectodermal defects-clefting syndrome
ANNOVA	A one-way analysis of variance
ANNOVAR	Functional annotation of genetic variants from high- throughput sequencing data
ARS	Axenfeld-Rieger syndrome
ARVC	Arrhythmogenic right ventricular cardiomyopathy
AT	Annealing temperature
AXIN	Axis inhibition protein
Barx1	Bar H-like homeobox gene1
BL	Bucco-Lingual
BMP	Bone morphogenetic protein
Bp	Base Pair
BWDD	British Baraitser-Winter Dysmorphology Database
c.	coding
C	Cytosine
Ca^{2+}	Calcium
cDNA	Complementary DNA
CF	Cystic fibrosis
CGI	Cooperative genetic interaction
CLPED	Cleft lip/palate-ectodermal dysplasia syndrome
dbSNP	Data base of single nucleotide polymorphism
Del	Deletion
dH_2O	Distilled water
DKK1	Dickkopf 1 homolog (Xenopus)
DID	Diastrophic dysplasia
DLX	Distal-less homeobox

DMSO	Dimethylsulphoxide
DNA	Deoxyribonucleic acid
dNTP	Deoxynucletriphosphate
DTDST	Diastrophic dysplasia sulfate transporter
ECM	Extracellular matrix
ED	Ectodermal dysplasia
EDA	Ectodysplasin
EDAR	Ectodysplasin receptor
EDTA	Ethylenediaminetetraacetic acid
EEC	Ectrodactyly-ectodermal dysplasia-clefting
EGF	Epidermal growth factor
EGFR	Epidermal growth factor receptor
FGF	Fibroblast growth factor
FGFR	Fibroblast growth factor receptor
FOXE1	Forkhead box E1 thyroid transcription factor 2
fx	Frameshift
Fz	Cell surface receptor Frizz
G	Guanine
GWAS	Genome-wide association studies
GLI	Glioma-associated oncongene homologue
Gsc	The homeobox goosecoid
h	Hour
HED	Hypohidrotic ectodermal dysplasia
HGF	Hepatocyte growth factor
Hh	Hedgehog
Hip	Hedgehog interacting protein
HPE	Holoprosencephaly
ins	Insertion
IP	Incontinentia pigmenti
IPH	Incisor-premolarhypodontia
IRF6	interferon regulatory factor-6
Islet1	ISL1 Transcription factor, Lim/homodomain
K	Keratin

Kb	Kilobase pairs
KLK10	Kalikrein10
KREMEN1	Kringle domain-containing transmembrane protein 1
LEF	Lymphoid enhancer-binding factors
LH	Lime/homeobox domain
M	Molar
Mb	Megabase pairs
MD	Mesio-Distal
Mg^{2+}	Magnesium
min	Minute
mM	Milli Molar
mm	Millimetre
mRNA	Messenger ribonucleic acid
MSX	Drosophila muscle segment homeobox
NCBI	National centre for biotechnology information
ng	Nanogram
NGS	Next generation sequencing
nM	Nano Molar
OD	Optical density
OODD	Odonto-onycho-dermal dysplasia
OFD	Oral-facial-digital syndrome
OMIM	Online Mendelian Inheritance in Man
P.	Protein
p	Short arm of chromosome
PAX	Paired box
PCR	Polymerase chain reaction
Pitx	Pituitary homeobox
PRKG1B	Protein kinase, cGMP-dependent, regulatory, type 1
Ptch	The multipass membrane protein, Patched
q	Long arm of chromosome
RNA	Ribonucleic acid
RT	Room temperature
Runx	Runt-related transcription factor

s	Second
Shh	Sonic hedgehog
SIX3	Sine oculis homeo box homolog3
SM	Steatocystoma multiplex
SMOC2	SPARC related modular calcium binding 2
SNP	Single nucleotide polymorphism
SNV	Single nucleotide variants
STRN3	Striatin, calmodulin binding protein 3
SVs	Structural variations
T	Thymine
TBE	Tris borate EDTA
TCF	T cell factor
TGIF	TGFB-induced factor homeobox
TGFß	Transforming growth factor ß family
T P63	Tumour protein P63
TNF	Tumour necrosis factor
TNS	(Witkop) Tooth-nails syndrome
UTR	Untranslated region
UV	Ultraviolet
V	Volt
v	Version
VWS	Van der Woude Syndrome
Wnt	Wingless
WNT10A	Wingless-type MMTV integration site family member 10A
ZNF	Zinc finger
ZFYVE26	Zinc finger, FYVE domain containing 26

Gene and protein symbol conventions:

Human gene: Italicised, with all letters in uppercase (e.g., *MSX1*).

Human protein: Not italicised, with only the first letter in uppercase (e.g., Msx1).

Mouse gene: Italicised, with only the first letter in uppercase (e.g., *Msx1*).

Mouse protein: Not italicised, with all letters in uppercase

Publications and abstracts arising from this thesis

Peer reviewed publication

S. Alfawaz, F. Fong, V. Plagnol, F.S.L. Wong, J. Fearne, D. P. Kelsell.

Recessive oligodontia linked to a homozygous loss-of-function mutation in the SMOC2 gene. Arch Oral Biol. 2013 May;58(5):462-6

Abstracts

S. Alfawaz, M. Hector, D. P Kelsell and F. S. L. Wong (September 2012) Crown Dimensions of the Permanent Teeth of Family with Hypodontia. *J Dent Res* 91(Spec Iss C): 313, 2012 (www.dentalresearch.org). Presented at the sixth International Association for Dental Research Pan-European Region Meeting (IADR/PER) in Helsinki, Finland. (Oral presentation)

S. Alfawaz, F.S.L. Wong , M. Hector and D. P .Kelsell (July 2011) Phenotypic and genotypic features of hypodontia in Saudi families. *Int J Paediatr Dent.*2011 Jun; 21(S1):1-272. Presented at the 23rd International Association of Paediatric Dentistry (IAPD) Congress in Athens Greece.

S. Alfawaz, F. S. L. Wong, M. Hector and D. P. Kelsell (June 2011) Phenotypic and genotypic features of inciso-premolar hypodontia in Saudi families. Presented at the 2nd WUN Symposium for Oral Health Sciences in Leeds, UK.

CHAPTER 1: INTRODUCTION

1.1 Background and General Aim

1.1.1 Background

Agenesis of one or more teeth is a common defect in man. It can be regarded as a spectrum of diseases with the mildest and most common phenotype being non-syndromic selective tooth agenesis (hypodontia). Third molars are the most commonly missing teeth in the dentition, followed by the second premolars and upper lateral incisors (Matalova *et al*., 2008). These tooth types are also the most commonly missing in the more severe phenotype (oligodontia).

Tooth development can be affected by either environmental or genetic factors. Children with malignant diseases at tooth-developing ages are at high risk of tooth agenesis from treatment. Irradiation therapy has more severe effects on tooth development than chemotherapy (Holtta *et al*., 2005a). Many genes have been found to affect tooth development through gene expression and experimental studies in mice. In theory, any of these genes may cause tooth agenesis (for a review see Nieminen (2009)). Family studies show that tooth agenesis, in an isolated form, can be caused by mutations in the *MSX1, PAX9, AXIN2, EDA, WNTA10* and *SMOC2* genes, with autosomal and X link traits of inheritance. As the phenotypes and genotypes of tooth agenesis are diverse, it is the aim of the study to determine the type of inheritance and phenotype, genotype of familial tooth agenesis among consanguineous families.

1.1.2 General aim

The general aim of this study was to investigate the spectrum of non-syndromic tooth agenesis in consanguineous families by characterising the clinical features of the phenotypes and to verify the genotypes of these phenotypes using linkage analysis and/or whole exome analysis.

1.2 Tooth agenesis

1.2.1 Definition, terminology and classifications

Tooth agenesis is a general term that has been used to describe developmental failure of teeth. It is defined as a congenital lack of one or more deciduous or permanent teeth that results from premature arrest of tooth development during the early stages of tooth development before the mineralisation stage (Thesleff, 2006). It has been classified as handicapping dento-facial anomalies by the World Health Organisation, defining that this anomaly can cause disfigurement, or impedes function, and may require treatment because it is likely to be an obstacle to the patient's physical or emotional wellbeing (World Health Organisation, 1962).

Like teeth eruption time which has a wide variation, the variation in a tooth's mineralisation and calcification stage between individuals is wide and is dependent on the gender, race and environmental factors. The diagnosis of tooth agenesis is usually detected clinically and confirmed by radiographs. Clinically, by the age of three all deciduous teeth should have erupted, and by the age 14 for all permanent teeth except the third molars. Radiographically all the deciduous teeth and the crypts of first permanent molars should be visible at birth. By the age of six, all permanent tooth crowns except the third molars should be visible. However second premolars sometimes give a false-positive diagnosis of tooth agenesis in radiographs as they may have late onset of mineralisation. Therefore, it has been suggested that diagnosis of tooth agenesis in permanent teeth shouldn't be made before the age of six, excluding third molars (Thesleff, 1995; Bailleul-Forestier *et al*., 2008). Third molars developments are much more varied in their formation; they usually start to mineralise by the age of 8 but may have a very late presentation in radiographs (Arte *et a*l., 2001).

Different terms have been used to describe the spectrum of tooth agenesis. These terms are mainly applied according to the number of teeth missing. Hypodontia and oligodontia are the most common terms used in the literature to name this type of dental anomaly. Hypodontia has been used generally to describe the congenital lack of tooth development, but most recently it has been used specifically for the most common types with only a few teeth missing (six or less excluding the third molars). Oligodontia is now generally used for more severe type of teeth agenesis, with more than six teeth missing excluding the third molar. According to Nieminen (2009), these terms might mislead researchers as the definition of oligodontia is arbitrary and does not show the severity of the condition since it does not include the third molars. He recommends the use of the terms partial, selective or severe tooth agenesis, as they directly refer to the developmental failure (Nieminen *et al.,* 2009). There are other terms that have been used in the literature to describe the reduction in number of teeth, such as lack of teeth, aplasia of teeth, absence of teeth, congenitally missing teeth and incisor-premolar hypodontia (IPH) which applies to incisor and premolar agenesis only. The term anodontia denotes the very rare extreme type with a complete absence of teeth (Nieminen *et al.,* 2009). In this thesis the term tooth agenesis was used as it has been used more recently, as an informative term directly linked with the underlying developmental defect.

The classification of tooth agenesis based on its severity (including third molars) is: partial or selective (equivalent to hypodontia), severe (oligodontia) and complete (anodontia). This condition can be manifested as an isolated or non-syndromic disorder, or as part of a syndrome. Isolated tooth agenesis can be familial or sporadic (Arte *et al.,* 2001).

1.2.2 Prevalence of tooth agenesis

Dental clinicians claim that tooth agenesis has increased during recent decades. Until now it is not clear whether this observation is related to an improvement in imaging over the years, increasing dental awareness, or whether it represents a real drift towards increased prevalence of tooth agenesis. This claim has been investigated by Mattheeuws *et al.*, (2004) and their meta-analysis concluded that available data are too limited to confirm a possible trend in human dentition, because a period of six decades is too short to investigate an evolutionary change in humans. However, the data confirmed that there was a sudden jump in the prevalence of missing teeth since 1965. These authors suggested two explanations for this phenomenon. The first explanation was that this increase in prevalence was an outcome of an improvement in dental examination facilities (imaging) and awareness about tooth agenesis, which means that the true prevalence might have remained unchanged, the increase was only in the reported prevalence (Mattheeuws *et al.*, 2004). The second explanation was that this rapid increase was due to an unidentified environmental factor affecting the phenotype as Brook (1984) suggested in a multifactorial model (Brook 1984; Mattheeuws *et al.*, 2004) (see Section 1.3.1.4).

1.2.2.1 Tooth agenesis in deciduous dentition

Tooth agenesis in deciduous dentition is less common than in permanent dentition, with a prevalence of less than 1% in the general population (Vastardis 2000; Arte *et al.*, 2001) with the most commonly affected tooth type being the incisor. The upper lateral deciduous incisors are the most commonly missing, accounted for 50% of the missing deciduous teeth, followed by the lower incisor; both together are accounted for 90% of all affected deciduous teeth (Grahnen and Granath 1961; Ravn 1971; Jarvinen and Lehtinen 1981). The prevalence of deciduous tooth agenesis seems to be more common in Japan as it has been reported to have a

high prevalence of 2.4%, with the most affected tooth being the lower lateral incisor (Yonezu *et al.,* 1997).

The agenesis of deciduous molars, canines or upper central incisors has rarely been reported as an isolated or non-syndromic case (Grahnen and Granath 1961; Ravn 1971; Daugaard-Jensen *et al.,* 1997). In Caucasian populations no significant difference exists in the prevalence of tooth agenesis between genders in the deciduous dentition (Nieminen, 2009).

The association between agenesis in deciduous and permanent dentitions has been confirmed (Whittington and Durward, 1996; Daugaard-Jensen *et al.,* 1997b). According to Matalova *et al.,* (2008) it is a rule that if the deciduous tooth does not develop, then the successor permanent tooth will be missing because permanent tooth buds off from the deciduous tooth bud (Matalova *et al.,* 2008). Syndromic tooth agenesis especially syndromes with ectodermal dysplasia, show tooth agenesis in the deciduous dentition as a phenotype (see Section 1.7.2).

1.2.2.2 Tooth agenesis in permanent dentition

1.2.2.2.1 Selective (partial) tooth agenesis

Tooth agenesis in permanent dentition is a prevalent dental anomaly in humans. Most of the reported cases of developmental failure of teeth were for one or a few teeth only. The severe phenotype of tooth agenesis is rare (Nieminen, 2009) with a prevalence of only 0.16% in a Danish study (Rolling and Poulsen, 2001). The prevalence of permanent selective tooth agenesis ranges from 1.6% to 9.6% in the general population excluding third molars, and it varies with tooth type (Vastardis, 2000). Including the third molars, the prevalence is higher, up to 25% of all reported cases (Matalova *et al.,* 2008). However, these studies on the prevalence of tooth agenesis were carried out on varied populations and ethnic origins. In Saudi Arabia where the population is mainly Arabs, two studies were carried out in Riyadh,

in the central region, to investigate the prevalence of selective tooth agenesis (hypodontia) and peg-shaped upper lateral incisors. The first study, by Al-Emran *et al.*, in 1990, examining 500 Saudi male students within the age range of 13-14 years, reported a prevalence of 4%. The most frequently affected tooth in this study was the lower second premolar, followed by the upper laterals, and upper second premolars. The peg-shaped upper lateral incisor was observed in 4% of the sample (Al-Emran, 1990). The other study was by Salama and Abdel-Megid in 1994. In this study the sample was higher, including 1300 Saudi children with an age range between 5-10 years. They found that selective tooth agenesis prevalence was only 2.6%, with the lower second premolar being the most commonly missing tooth. Peg-shaped permanent upper lateral incisor was reported to be in 0.7 % of the sample. The tooth agenesis was almost equally distributed between upper (52%) and lower (48%) arches (Salama and Abdel-Megid, 1994). Since the earlier study was in an older age group, it is expected that they would find late developed premolars. However, the prevalence was higher than the latter study. Also, from the age groups and years of study, there should not be any overlap in the study subjects. Hence, a possible explanation for the difference in tooth agenesis prevalence between these two studies might be due to the difference in the sample sizes, unless there was an ethnic shift in the latter years.

In African Americans, agenesis prevalence was 7.7% with the lower second premolar as the mostly affected (Jorgenson, 1980). In Japan the prevalence was estimated to be 9.4% and the most frequently missing tooth was the lower second premolar. Interestingly in this study, the lower lateral incisor was also frequently affected, accounted for 18.8% of missing teeth, in contrast to the prevalence in other populations (Goya *et al.*, 2008). Agenesis of the permanent

lower incisor has been observed as a characteristic dental abnormality of individuals of Asian ethnicity, mainly Japanese and Chinese (Niswander and Sujaku, 1963; Davis, 1987).

Polder *et al.*, (2004) in a meta-analysis reviewed tooth agenesis reports and found that in Caucasian populations, the most commonly missing tooth is the lower second premolar, followed by the upper lateral incisor and upper second premolar, comprising about 85% of all missing teeth. The most stable permanent teeth were the upper central incisors (0.016%) and lower first molars and canines (0.03%). Polder *et al., (*2004) also reported that tooth agenesis pattern is often bilateral (Polder *et al.*, 2004). Bailit (1975) has also noticed that in selective tooth agenesis the most common pattern was bilateral and symmetric, with the exception of cases where the upper lateral incisors at the left side were more affected than the right side (Bailit, 1975). Lundstrom, by contrast, has observed that agenesis of teeth was mostly unilateral; about half of the missing teeth were absent unilaterally (Lundstrom, 1960). It has been suggested that peg lateral incisors, rudimentary third molars, and unilateral tooth agenesis are variable expressions of the same mutated gene of tooth agenesis, and might reflect the incomplete penetrance of the causative mutation (Woolf, 1971; Stritzel *et al.*, 1990; Vastardis, 2000).

Interestingly, some researchers noticed a connection between the number of missing teeth and the type of tooth affected. For example, Muller *et al.*, (1970) determined that if the missing teeth are two or less, upper lateral incisor is the commonest tooth type to be affected, but if the number of missing teeth is three or more, the second premolar is the commonest type (Muller *et al.*, 1970). A study focusing on second premolars agenesis showed, by contrast, that the absence of one second premolar is the most common phenotype and the absence of three premolars is the least common phenotype (Stritzel *et al.*, 1990).

Gender predominance in tooth agenesis has been investigated thoroughly, and a number of reports found no statistically significant sex predominance of tooth agenesis (Muller *et al.*, 1970; Haavikko, 1971; Thompson and Popovich, 1974; Magnusson, 1977; Rolling, 1980; Davis, 1987). However, some reported a female to male ratio of between 3:2 and 6.5:3.5 (Egermark-Eriksson and Lind, 1971; Brook, 1974; Bergstrom, 1977).

1.2.2.2.2 ***Severe tooth agenesis (oligodontia)***

Severe tooth agenesis is a rare congenital defect. In European populations, it was seen at a level of around 0.25% (Sarnas and Rune, 1983). The prevalence was found to be 0.08% in a Dutch study (Schalk-van der Weide *et al*., 1992), and 0.16% in a Danish study (Rolling and Poulsen, 2001). In Japan, the prevalence was found to be slightly higher, about 1.4% (Goya *et al*., 2008).

The gender difference of oligodontia between males and females was not found to be statistically significant neither was the difference in distribution of missing teeth between the maxilla/mandible and left/right sides (Schalk-van der Weide *et al.*, 1993); (Rolling and Poulsen, 2001). The most affected teeth are the second premolars and/or upper lateral incisors. It has been reported that in oligodontia, two of every three congenitally missing teeth are second premolars and/or upper lateral incisors (Rolling and Poulsen, 2001).

1.2.2.2.3 ***Anodontia***

Anodontia (congenital lack of all teeth) without associated abnormalities is extremely rare and occurs as an extreme dental phenotype in ectodermal dysplasia syndrome (Arte *et al.*, 2001; Nieminen, 2009).

1.3 Tooth agenesis aetiology

Tooth agenesis is an interesting complex disorder, and in recent decades several theories have been suggested to explain it and verify its aetiology. Currently the aetiology of tooth agenesis is related generally to the role of genetic and environmental factors (Nieminen, 2009).

1.3.1 Theories about tooth agenesis

Tooth agenesis as a dental defect has been an active area of research; in the last century scientists tried to apply evolutionary and anatomic models to the developmental failure of teeth, for example, odontogenic polarity, Butler's field theory, the anatomical model and Sofaer's model of compensatory tooth size interactions (Vastardis, 2000). The theories are described in this thesis in chronological order to show the attempt that has been made during these years to investigate this condition.

1.3.1.1 Butler's theory (1939)

This theory tried to explain the observation of developmental failure of certain teeth more than others, by dividing the mammalian dentition into three morphologic regions, incisors, canines, and premolars/molars. Within these regions, one **"**key**"** tooth is supposed to be stable in each quadrant. The morphologically related teeth within the region become gradually less stable the further they are away from the 'key tooth'. So, the key tooth in the molar/premolar area would be the first molar, and the second and third molars at the distal end of the region, and the first and second premolars at the mesial end would be less stable. According to this theory, the third molar and the first premolar would be expected to be more variable dimensionally and morphologically.

Epidemiological studies confirm this observation for the third molar, but not for the first premolar. Buckley and Doran (2001), reported an interesting case of a family with non-

syndromic oligodontia, a pair of monozygous twins and their mother with patterns of tooth agenesis different to what Butler has expected, each twins had 13 missing permanent teeth and their mother 16 missing teeth. The twins had unerrupted third molars and all three family members were showing upper lateral incisors (Buckley and Doran, 2001). Another case, reported recently by Hosur *et al*., (2011) of a 12-year-old boy with non-syndromic oligodontia, who was missing all his permanent teeth except premolars (the diagnosis of the third molars was not confirmed yet) (Hosur, 2011). However, Bailit (1975) has tried to explain this contradiction with regard to the premolars by assuming that the earliest mammals had four premolars, whereas some higher primates, including man, have lost the first two. These missing teeth would have been farthest from the key tooth and in an evolutionary sense could be considered unstable (Bailit, 1975).

1.3.1.2 Clayton (1956)

Clayton (1956) attempted to prove Butler's theory by suggesting that the most common missing teeth were vestigial organs and have no value for modern man in the evolutionary process. He examined 3557 subjects and noticed that the most posterior tooth of a tooth group (incisors, premolars and molars) was the most frequently missing in the sample (Clayton, 1956). Some scientists believe that future man's dentition will only have one incisor, one canine, one premolar and two molars for each quadrant (Coon, 1962; Pindborg, 1970). This hypothesis needs evidence of a link of more favourable improvement in health with the reduction in teeth number.

1.3.1.3 Compensatory tooth size interactions model (1971)

Sofaer *et al.*, (1971) suggested that "if a tooth which developed early was larger than the average size, then its later developing adjacent tooth would have a higher chance to be smaller than normal, and vice versa" (Sofaer *et al.*, 1971). According to these authors, the reduction in tooth size and tooth agenesis is a compensatory interaction between tooth germs

during development (*e.g.* lateral incisors develop after the centrals and their initiation depends on the availability of essential local needs. The absence or reduction in size of the teeth on one side induces a compensatory increase in size of the teeth of the contra-lateral side (Sofaer *et al.*, 1971).

1.3.1.4 Multi factorial model of dental anomalies of tooth number and size (1974)

Brook, in his hypothesis, proposed a model (Figure 1.1) where genetic factors play a major part in dental anomalies of tooth number and size, with an influence also from environmental factors (Brook, 1974; Bailit, 1975; Brook, 1984). This model includes two separate curves for males and females, the two tails of the curves demonstrate the relationship between different dental anomalies: tooth agenesis and microdontia lie at one end and supernumerary and macrodontia at the other end. Bailit (1975) suggested a polygenic model and agreed with Brook in his hypothesised model; both suggested that tooth agenesis illustrates the quasi-continuous trait. According to both, the variable expression of tooth agenesis reports especially in monozygotic twins, supports the hypothesis of this model (Brook, 1974; Bailit, 1975; Brook, 1984).

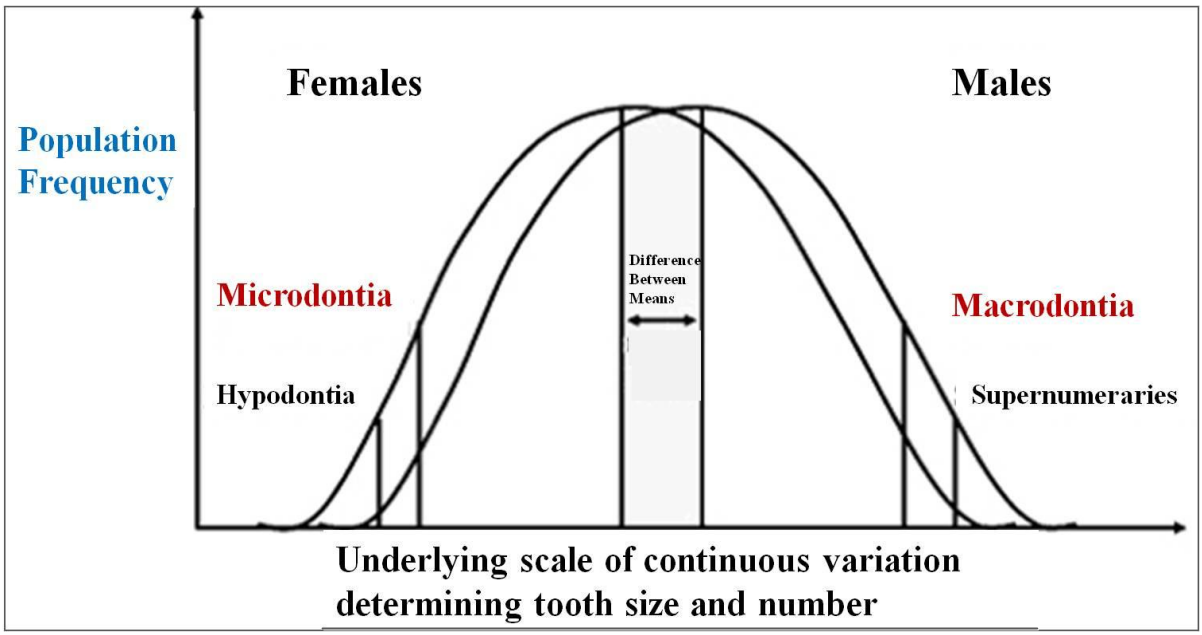


Figure 1.1: The multi-factorial model hypothesised by Brook (Brook, 1984).

1.3.1.5 Anatomical model (1988)

Svinhufvud and his co-workers (1988) suggested an anatomical model as an alternative to the evolutionary model to explain the selectivity of tooth agenesis. They argued that during tooth development (*e.g.* areas of embryonic fusion), certain regions are more prone to epigenetic influences and therefore agenesis. They listed some regions and called them fragile areas. The first was the junction between the pre-maxilla and maxilla. This can be applied to maxillary incisor instability, but not for the significant stability of the canine which is also next to that junction. The second region was in the mandible and they called it fragile because it is at the distal end of the primary dental lamina. However, this region of mandibular agenesis seems to be specific for permanent dentition but not for the deciduous teeth as the absence of second deciduous molars is rare. The third region was the fusion of the two mandibular processes where agenesis of lower central incisors occurs commonly (Svinhufvud *et al.*, 1988).

1.3.1.6 Neural developmental fields model (1997)

In this model Kjaer (1997) related the function of peripheral nerves to tooth agenesis. He proposed neural developmental fields in the maxilla and mandible (incisor field, canine/premolar, and molar field) as an explanation for the common sites of tooth agenesis. His hypothesis was based on the lack of innervation in the location of tooth agenesis. The area where the innervation ends in any of these fields (incisor field, canine/premolar, and molar field) is more likely to manifest tooth developmental failure (Kjaer, 1997).

The above models suggested different explanations for tooth agenesis phenomena. However, though they could apply to some mild phenotype, they failed to apply to the severe phenotype and in tooth agenesis phenotypes with variable penetrance in which one side was affected more than the other among the affected members.

1.3.2 Environmental factors

The sensitivity of tooth development to environmental disturbances has been shown in many studies. Many environmental factors might affect and arrest tooth development such as traumatic dental injuries, cancer therapy and dioxin. The direct relation between tooth development aberration and childhood radiotherapy (Holtta *et al.*, 2005a), chemotherapy and stem cell transplantation has been confirmed (Holtta *et al.*, 2005a). The dental defects may include tooth agenesis, microdontia, developmental delays, root anomalies and calcification defects (Holtta *et al.*, 2005a, 2005b). The severity of defects and number of teeth involved in this environmental insult are influenced by the age of the child when he/she received the treatment (Pedersen *et al.*, 2012).

The effect of dioxin on tooth development has been reported as an increase in prevalence of hypomineralisation of teeth in Kotka and Anjalankoski in Finland (Holtta *et al.*, 2001) and as an increase in prevalence of tooth agenesis and hypomineralised enamel among people exposed to it in Seveso, Italy (Alaluusua *et al.*, 2004). This has been verified in the laboratory in experimental animals (Lukinmaa *et al.*, 2001; Yasuda *et al*., 2005). Bailit believed that prenatal environmental factors such as nutrition and diseases have more influence than the postnatal factors on dental development (Bailit, 1975). No specific aetiological relationship has been found between tooth and systemic diseases or endocrine disturbances (for review see Grahnen, 1956; Schalk-van der Weide *et al.*, 1992). This thesis focuses on genetic factors as an explanation of tooth agenesis in the cases reported.

1.3.3 Genetic factors

A key role for genetic variation in tooth agenesis has been demonstrated by familial studies, twin reports and association with inherited syndromes (Wong *et al*., 2000; Nieminen, 2009), and confirmed by identified gene defects (Vastardis *et al.* 1996; Stockton *et al.*, 2000;

Nieminen *et al.,* 2001). The molecular basis of tooth agenesis is not fully understood, but odontogenesis studies in both humans and in mouse model show that tooth development is strictly directed by genes, these genes determine the position, number, size and shape of teeth. For example several genes in tooth development have been identified from mutant knockout mice models (Kollar and Baird, 1969, 1970; Lumsden, 1988; Mina and Kollar, 1987; Sharpe, 2001). More than 300 genes have been identified in developing tooth germs, P. Nieminen supported by the University of Helsinki has created a graphical data base that shows the expression patterns of these genes in schematic views (http://bite-it.helsinki.fi). Most of the genes in this database have roles in cellular communications such as Bone Morphogenetic Proteins BMPs, Fibroblast Growth factors FGFs, Sonic hedgehog Shh, Wingless signalling pathway Wnt and transcription factors. Mutations in a number of these genes have arrested tooth development in mice and man (Thesleff, 2006) .

1.3.4 Polygenic factors

The multifactorial or polygenic model has been proposed, as the tooth agenesis phenotype is very variable, which means that it is not entirely due to a single gene defect. This observation has been evidenced in monozygotic twins and triplets (Boruchov and Green, 1971; Moller *et al*., 1981; Brook, 1984; Townsend *et al.,* 1995). There is no dominance or recessivity in the polygenic model. Many assumptions have been suggested to simplify this model. For example, each contributing gene has equal effects that are too small to pass a stringent genome-wide significance level, but their additive effects result in phenotype expression. As these genes interact as an incomplete dominance model, there is no linkage involved and large sample sizes are required, thus many collaborations have been established to prove this concept. Genome-wide association studies (GWAS) have provided a great attempt to discover hundreds of genetic variants that are involved in complex diseases and traits such as

rheumatoid arthritis and cardiovascular disease (Simonson *et al.*,2011). Two statistical analyses have been used to assess the involvement of common SNPs that are below the genome-wide significance level: polygenic analysis and mixed linear modelling, these methods test many SNPs in aggregate for collective effect on phenotype (Stahl *et al.*, 2012).

1.4 Tooth Development

Mammalian teeth develop from the interaction of the oral epithelial cells and the underlying mesenchymal cells. Human dentitions consist of 20 deciduous and 32 permanent teeth. The basic developmental process is similar for all teeth. This process involves characteristic bud, cap, and bell stages, but each tooth develops anatomically as a distinct element in its tooth group.

1.4.1 Phases of tooth development

A tooth develops in a continuous process characterised by a sequence of phases defined by their shapes: bud, cap and bell phases (Figure 1.2). The initial phase is the bud and is distinguished histologically in mice as a rounded, localised growth of epithelial cells bounded by proliferating mesenchymal cells around the 11^{th} to 13^{th} day of embryonic development (E11-13). Then the rounded bud epithelial cells increase and expand to form a concave shape which is the beginning of the cap stage at E14. The epithelial cells will form the enamel organ and mesenchymal cells will form the dental papilla. The tissue surrounding these two structures is the dental follicle. After additional cell differentiation of the enamel organ and dental papilla, the tooth reaches the bell phase at E16, which is characterised by morphological and histological differentiation. This process is controlled by the enamel knot which is a signalling centre that forms at the tip of the late bud (Figure 1.2). By the late bell stage at E18, ameloblasts derived from epithelium form the enamel, and the dental papilla gives rise to the pulp, dentine and periodontum (Avery and Chiego, 2006).

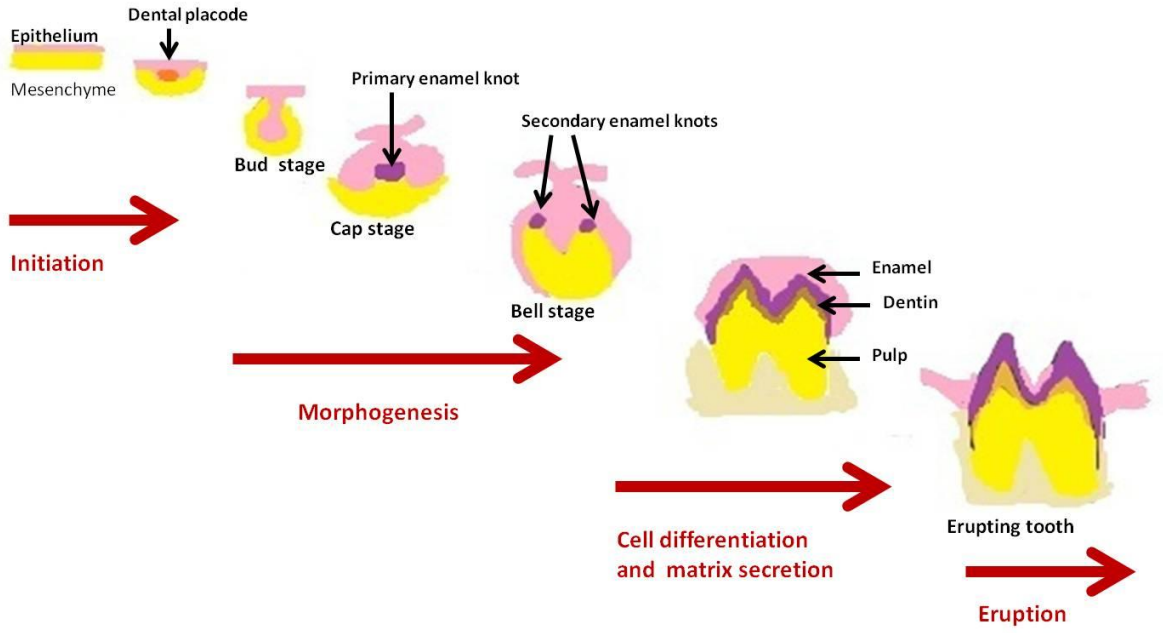


Figure 1.2: Schematic presentation of the morphology of tooth development adapted from Nanci and Ten Cate (2008).

1.4.2 Signalling pathways and genes expressed during tooth development

Signalling molecules between cells consist of several different families of growth factors and their specific receptors. The families most studied and reported are the transforming growth factor ß family (TGFß), which includes fibroblast growth factor (FGF), bone morphogenetic proteins (BMPs), epidermal growth factor (EGF), as well as the hedgehog (Hh) and the wingless (Wnt). They function reciprocally between the epithelium and mesenchyme during tooth morphogenesis. It is evident that the localised expression of signals is controlled by epithelial-mesenchymal interactions. Interestingly, these signals are restricted to either epithelial or mesenchymal tissues or to a specific developmental phase (Thesleff and Sharpe,

1997). However, some of these signals are used repeatedly during different phases of tooth morphogenesis (Thesleff, 2000).

1.4.2.1 FGF signalling pathway

Several members of the highly conserved fibroblast growth factor (FGF) family have been involved in tooth development. The first to be identified was *Fgf3 (int-2).* Its expression is restricted to the dental papilla mesenchyme, while the *Fgf4*, *Fgf8* and *Fgf9* are expressed only in dental epithelial cells (Thesleff and Aberg, 1997). *Fgf10* is expressed lately in the mesenchyme, however its ability to stimulate cell proliferation is limited to the dental epithelium but not in the mesenchyme (Zhang *et al.,* 2005). However functional redundancy for FGF3 and FGF10 in tooth development has been proposed, as *Fgf3-l-or Fgf10-l-* mice Show minimal effect on tooth development while elimination of the *Fgf3* and *Fgf10* receptor (FgfR2b) arrests tooth development at the bud stage. *Fgf8* is expressed in the presumptive dental epithelium, mainly in tooth initiation before the budding phase. While *Fgf4* presents only in the enamel knots, whereas *Fgf9* is expressed widely in the dental epithelium to cover the domains of both *Fgf4*, *Fgf8* and the final differentiation of the odontoblasts and ameloblasts during the bell stage (Thesleff and Sharpe, 1997; Zhang *et al.,* 2005), (see Section 1.4.5 and1.4.6).

1.4.2.2 BMP signalling pathway

The role of bone morphogenetic proteins (BMPs) in tooth development was first identified by Vainio *et al*., (1993), who studied Bmp4 expression in the transmission of inductive interactions between the dental epithelium and mesenchyme (Vainio *et al*., 1993). *Bmp2*, *Bmp4* and *Bmp7* are expressed early in the dental epithelium, but later when the odontogenic potential shifts from the epithelium, *Bmp4* is expressed in the mesenchyme (Thesleff and Sharpe, 1997). However the contribution of these signals cannot be directly assessed as targeted mutations in *Bmp2* and *Bmp4* result in embryonic lethality before tooth germ

development (Winnier *et al.*, 1995). Nonetheless, analysis of the effect of mutations of some homeobox genes has revealed some details on the significant role of BMP signalling in early odontogenesis (Thesleff and Sharpe, 1997; Zhang *et al.*, 2005; Cudney and Vieira, 2012), (see Section 1.4.5 and1.4.6).

.

1.4.2.3 Hedgehog signalling pathway

The sonic hedgehog (Shh) is one of the main signalling pathways that involved in growth and patterning of a number of organs, including the tooth signalling pathway. It is expressed in dental epithelium in the early epithelial thickenings and in the enamel knot (Thesleff, 2006). Interestingly, its receptor, the multipass membrane protein Patched (*Ptch*), is expressed widely in the dental mesenchyme and in the epithelium area not expressing *Shh. Ptch* and a recently identified binding protein, hedgehog interacting protein (*Hip*) are transcriptional targets which negatively modulate hedgehog signals (Couve-Privat *et al.*, 2004). Glioma-associated oncogene homologue is a family of zinc finger transcription factors that includes *Gli1, Gli2 and Gli3* in vertebrates. This family mediates the transcriptional effects of *Shh* pathway. *Gli* homozygous mutations in mice (*Gli2-/-* and *Gli3-/-*) result in lack of tooth buds (Zhang *et al*., 2005; Nanci and Ten Cate, 2008).

Ptch and *Gli1*gene expression in the epithelium and the mesenchyme of the tooth suggests that *Shh* might signal directly to both epithelial and mesenchymal components (Dassule and McMahon, 1998). Up to now the role of *Shh* signalling in early epithelium thickening has been unclear, but it follows the expressions of *Fgf8* and *Fgf9*, which means that they might play a role in regulation of *Shh* signalling (Thesleff and Sharpe, 1997). Dassule *et al.*, demonstrated that *Shh* is required for normal growth, morphogenesis and polarity within the

tooth, but not essential for differentiation of the tooth (Dassule *et al.,* 2000) (see Section 1.4.5 and1.4.6).

1.4.2.4 Wingless signalling pathway Wnt

Wnt signalling consists of a network of proteins that participate in fundamental biological processes in embryogenesis, such as cell proliferation, migration, polarity establishment and stem cell self-renewal. Deregulation of Wnt signalling is also associated with cancer development. Wingless protein binds to its cell surface receptor Frizz (*Fz*) and activates an intracellular signalling pathway containing the armadillo protein (*β-catenin*). Binding of *β-catenin* with the transcription factors of the lymphoid enhancer-binding factor (*LEF*) /T cell factor (*TCF*) family switches the inhibitory state of LEF/TCF to an activating state and turns on the transcription of Wnt target genes (Gao and Chen, 2010). Lymphoid enhancer-binding factors are a group of transcription factors which bind to DNA through a high mobility domain, and are members of a family of nuclear proteins that includes the T-cell factor proteins, TCF. *Lef1* is first expressed in dental epithelial thickenings, but during bud formation shifts to be expressed in the condensing mesenchyme. Tooth development is arrested at bud stage in *Lef*1 knockout mice (Zhang *et al*., 2005; Nanci and Ten Cate, 2008).

Several *Wnt* genes are expressed during tooth development, including *Wnt10* genes which are expressed in the early dental epithelium (Thesleff and Sharpe, 1997), (see Section 1.4.5 and1.4.6). It is also important to notice that AXIN2 expression induced by *Wnt* signalling produces a protein that regulates a negative feedback mechanism on the signalling mechanism itself (Chen *et al.,* 2009). The protein product of AXIN2 is a negative regulator of the canonical Wnt pathway that suppresses signal transduction by promoting the degradation of ß-catenin (Dao *et al.,* 2007). AXIN2 functions as a scaffold protein which is an important

component of a protein complex required for the phosphorylation of ß-catenin (Dao *et al.*, 2010).

1.4.2.5 Ectodysplasin gene EDA

Ectodysplasin genes belong to a family of signalling molecules of tumour necrosis factor (TNF). The ectodysplasin pathway contains the ligand ectodysplasin A1 (Eda-A1) the receptor Edar and the adapter molecule Edaradd (Li *et al.*, 2008). In human mutations of any of those genes lead to syndromic tooth agenesis (hypohydrotic ectodermal dysplasia HED) (Zhang *et al*., 2009; Wu *et al*., 2012) or to non- syndromic severe tooth agenesis (Tao *et al*., 2006). It has been reported that *Eda* regulates the activity of enamel knots through its receptor and its intracellular transmission signal *Edaradd* which is expressed in the knot cells (Harjunmaa *et al*., 2012). It also expresses around the developing crown from the cells of the external epithelium. Disruption of the *Eda*, its receptor *Edar*, and *Edaradd* in mice arrests tooth development (Tucker and Sharpe, 2004).

1.4.2.6 Paired box genes PAX

Paired box homeotic genes are a family of transcriptions factors that are characterised by the DNA-binding paired domain, which is encoded by the paired box. This is a conserved DNA region originally identified in *Drosophila* (Bopp *et al*., 1986). So far, nine different *PAX* genes have been identified in mammals, they fall into four different subgroups; pax group 1 (*PAX1* and *9*) pax group 2 (*PAX2, 5 and 8*), pax group 3 (*PAX3* and *7*) and pax group 4 (*PAX4* and *6*) (Bopp *et al*., 1986). These genes carry out essential functions during embryogenesis. Paired box gene 9 *PAX9* gene **(**OMIM 167416) plays an important role in tooth development (Thesleff, 2006). The expression of the *Pax9* gene in mice can be identified in the mandibular arch mesenchyme before any signs of odontogenesis, and the high gene expression levels are maintained until the late bell stage, after that the gene's production is reduced (Tucker and Sharpe, 2004), (see Section 1.4.5 and1.4.6). Mutant mice

with a homozygous mutation in *Pax9* gene (*Pax9-/-*) lack all their teeth as their development is arrested at the bud stage and they have several craniofacial abnormalities, whereas the heterozygous (*Pax9+/-*) mutant mice develop normally (Peters *et al.*, 1998). Interestingly in human heterozygous mutation may lead to severe tooth agenesis (Boeira and Echeverrigaray, 2013).

1.4.2.7 Drosophila muscle segment homeobox genes MSX

The drosophila muscle segment homeobox gene family encodes transcription factors that can control the expression of other genes; these genes code for a sixty amino acid protein (homeodomain) able to bind to a specific DNA sequence. They consist of 3 mammalian genes (*MSX1, MSX2* and *MSX3*), first cloned from mice and identified as homologous to the Drosophila muscle segment homeobox gene (msh). Later *MSX* genes have been isolated from a variety of species (Alappat *et al.*, 2003). *msx*1 is widely expressed in many organs especially at sites where epithelial-mesenchymal interaction takes place, such as in odontogenesis during the bud and cap stages of tooth development (Diamond *et al.*, 2006). This gene has been investigated thoroughly and many reports have been published about its role in tooth development. The MSX1 protein inhibits transcription of the target genes through its interaction with other transcription factors, such as DLX2 and 5, LHX2, PAX3 and PAX9 (Galluccio *et al.*, 2012) Homozygous *Msx1* mutant mice (*Msx1-/-*) lack development of the incisors and molar development is arrested at the bud stage. They also have secondary cleft palate and alveolar process under-development (Satokata and Maas, 1994), (see Section 1.4.5 and1.4.6).

1.4.2.8 Distal-less homeobox genes DLX

The distal-less homeobox genes belong to a family of homeodomain transcription factors. There are six known genes each in mice and humans, they are found in three convergently transcribed pairs. Each pairs is linked to a Hox cluster. In mice and humans, *Dlx1* and *Dlx2*

are linked to *Hoxd, Dlx3* and *Dlx4* (also known as *Dlx7* and *Dlx8*) are linked to *Hoxb,* and *Dlx5* and *Dlx6* are linked to *Hoxa* (Panganiban and Rubenstein, 2002). *Dlx1* and *Dlx2* are expressed in the neural-crest-derived mesenchyme of the proximal regions of the mandibular and maxillary process (Diamond *et al.*, 2006). They are located in the same chromosome (chromosome 2), whereas *Dlx5* and *Dlx6* are located in chromosome 7. *Dlx2* is also expressed in the epithelium of the distal aspects of the first branchial arch prior to the initiation of tooth formation. *Dlx1, Dlx2, Dlx5* and *Dlx6* are all expressed in the ectomesenchyme of the future lower molar sites, but only *Dlx1* and *Dlx2* are expressed in the ectomesenchyme of the future upper molar sites (Thomas *et al.*, 1997; Tucker and Sharpe, 2004) (see Section 1.4.5 and1.4.6).

1.4.2.9 Pituitary homeobox genes Pitx

Pituitary homeobox genes comprise a family of paired-like homeobox transcription factors, expressed in the brain, heart, pituitary, mandibular and maxillary regions, oral epithelium, eye and umbilicus. *Pitx2* gene acts as a transcriptional activator but it might act as a repressor by interaction with other transcription factors. *Pitx2* expression in oral epithelia enhances the expression of the *Fgf8* (Tucker and Sharpe, 2004). In *Pitx2* knockout mice, the expression of *Fgf8* is reduced and tooth formation is arrested at bud stage (Lin *et al*., 1999), (see Section 1.4.).

1.4.2.10 Tumour protein P63 TP63

The *P63* gene is a homologue of tumour-suppressor *P53* which is a transcription factor that regulates the progression of the cell through its cycle. *P63* gene is highly expressed in the basal layers of epithelial tissue and has two isoforms TAp63 isoforms and ΔNp63 isoforms (Rufini *et al.*, 2011). The role of the *P*63 in tooth development has been investigated thoroughly in a mouse model. *P*63 is expressed in the epithelium of dental lamina E10 and its expression extend to the bud and cap stages whereas in the bell stage of molar development,

P63 is evident in the outer enamel epithelium (Mills *et al*., 1999; Yang *et al*., 1999; Laurikkala *et al*.; 2006; Rufini *et al*. 2006). P63 regulates expression of some targets involved in tooth morphogenesis such as FGF signalling (FGF3, FGF6 , FGF7 and FGF10) (Candi *et al*., 2007), and BMP7, whose expression in mouse oral epithelium is vanished upon deletion of *P63* (Zouvelou *et al*., 2009). Lack of *P63* in mice arrests tooth development at an early stage, and affects developments of limbs, hair follicles and mammary glands (Mills *et al*., 1999; Yang *et al.,* 1999). Although reports linking *P63* to tooth development have been obtained from mouse models, there is evidence that it has the same role in human (Rufini *et al*., 2011). Several mutations have been identified in Hay-Well syndrome (McGrath *et al*., 2001), and in Ectrodatyly-ectodermal dysplasia-clefting (Wei *et al*., 2012).

1.4.3 Dentition patterning

Patterning refers to the process of determination of specific tooth types at their correct positions in the dental arches. In mammals, teeth are heterodent and consist of different forms in three groups: incisiform, caniniform, and molariform. Three main hypothetical models have been proposed to explain the patterning process. The first is the field model which is part of Butler's field theory (Section 1.3.1.1); this model predicts that all teeth primordia are initially equal, and for each tooth group the tooth shape determining factors exist in the ectomesenchyme in different concentrations and overlapping fields (Tucker and Sharpe, 2004).

The second is the clone model proposed by Osborn in 1987, which suggests that each tooth type is derived from a single clone of ectomesenchymal cells. These cells are non-equivalent for each of the tooth groups and programmed by epithelium to form teeth of a given pattern (Mitsiadis and Smith, 2006). These two models were stimulated from observation of human dentition, the shape of the teeth and their positions. The third model is called the dental homeobox code model. It has been shown that tooth identity can be changed by manipulation

of the epithelial signals and alteration of homeobox gene expression in the mesenchyme. (Townsend *et al*., 2009). This model is considered further below.

1.4.3.1 Homeobox code model

Tooth difference fields are established in the developing face from signals from the oral epithelium. The homeobox code model is based on observations of the expression of several homeobox genes in spatially restricted regions in the first branchial arch in mice before day 11 of embryogenesis. It has been noticed that early expression of *Msx1, Msx2* and *Isl1* homeobox genes prior to the initiation of the tooth germ is restricted to the anterior regions of the first branchial arch (distal, midline, ectomesenchyme) where incisors will form but not multi-cuspid teeth. (Mitsiadis *et al.,* 2003; Tucker and Sharpe, 2004). By contrast, expression of several of the *Dlx, Barx* and *Pitx* genes is restricted to the posterior regions of the first branchial arch where the future molars but not incisors will develop. *Barx* expression overlaps with *Dlx1*and *Dlx2* and in mice corresponds directly with ectomesenchymal cells that will grow into molars (Tucker and Sharpe, 2004). As the homeobox code model proposed that overlapping domains of these genes give the positional information for tooth morphogenesis, much evidence has been gathered from mouse experiments to underpin this model. In *Msx1*-|-double knockout mice, the incisors fail to form and molar formation is arrested at the late bud stage, while in the *Dlx1*-|-and *Dlx2*-|-mice only development of the upper molar teeth is arrested at the epithelial thickening phase. The lower molars develop normally because of the compensation action of other *Dlx* genes such as *Dlx-5* and *Dlx-6* that are only expressed in the ectomesenchyme of the mandibular primordium. Incisors form normally in these transgenic mice (Tucker and Sharpe, 2004; Nanci and Ten Cate, 2008; Townsend *et al.,* 2009).

1.4.3.2 Cooperative genetic interaction (CGI) model

This is a genetic developmental model that has been suggested lately by Mitsiadis and Smith (2006) to include all the factors that affect tooth patterning. It combines the clone and homeobox code models. In the CGI model, neural crest-derived cells, homeobox-containing genes and signalling molecules are all important in tooth patterning (Mitsiadis and Smith, 2006). However there are some phenotypes that have not been completely explained yet by these models.

1.4.4 Dental lamina formation, the first step of tooth development

During day 37 of embryogenesis, condensed horseshoe shaped epithelial bands form around the presumptive upper and lower jaws in the first branchial arch. These bands are the future dental arches of the two jaws. Each epithelial band (also known as the primary epithelial bands) separates into two parts, the dental lamina which forms first, and the vestibular lamina which develops later in front of the dental lamina to form the vestibule between the cheek and the tooth bearing area. The initiating dental lamina starts to function in the sixth week of prenatal development and continues to function until all 52 teeth are formed. The last teeth to develop are the third molars. From the leading edge of the lamina, 20 deciduous teeth develop, then the lamina continues to grow to form the permanent teeth which succeed the deciduous teeth; this part of the lamina is called the successional lamina. The general lamina then extends posteriorly along the jaw to form the 12 permanent molars (Nanci and Ten Cate, 2008)..

1.4.5 Initiation of the tooth and placode formation

The first sign of tooth development is the formation of ectodermal (dental) placodes on the oral surface of the first branchial arch at day 11 of embryogenesis, which are localised thickenings in the primary epithelial bands at the position of each tooth group (Figure 1.3). These placodes are the key features of initiation of the tooth (Thesleff, 2006). Interestingly

the same placodes begin the formation of all ectodermal organs such as hair, nails and mammary, salivary and sweat glands. Genes controlling the formation and function of these placodes are similar (Pispa and Thesleff, 2003).

1.4.5.1 Tooth initiation in the mouse model

Laboratory mice have been used to study the basis of tooth development, most of current knowledge about the biology of tooth formation is based on mouse, but it is important to consider the significant differences that do exist between the mouse and human dentition, *e.g.* mouse have only a single dentition during their live time, their dentition lacks both canines and premolars tooth type, and has one central incisor in each quadrant (Tucker and Sharpe, 2004).

The oral epithelium of mouse embryos can be divided in to two fields: the lateral field which characterised by the expression of the *Fgf8* and *Fgf9* overlying the future molars area and *Bmp4* which expressed distally overlying the future incisor area. These signalling molecules have been identified for their roles on homeobox genes (transcription factors) in the ectomesenchyme such as *Fgf8* which induces the expression of *Barx1* and *Dlx2*, whereas *Bmp4* suppresses the *Barx1* expression and induces the expression of *Msx1*and *Msx2*. The strict control of these signals as explained before in tooth patterning (Section 1.4.4) results in restriction of the *Barx1* and *Dlx2* to the presumptive molar area and *Msx1*and *Msx2* to the presumptive incisor area (Tucker and Sharpe, 2004). The expression of these homeobox genes can be manipulated in experiments, the inhibition of *Bmp* signaling early in mandible development, changed the presumptive incisor region in to molar region by inducing the *Barx1* expression leading to transformation of tooth identity from incisor to molar (Tucker *et al.*, 1998).

The transcription factors Lim-homeobox domain genes *Lhx6* and *Lhx7*, are the earliest markers in the neural crest-derived ectomesenchyme that have been identified so far, they expressed at day 9 of embryo development in mice (E9) (Grigoriou *et al*., 1998).

Dental recombination studies have investigated the results of separating and recombination of the epithelium and ectomesenchyme from the first and second branchial arches. The recombination of second arch mesenchyme with first branchial arch oral epithelium induced *Lhx6* and *Lhx7* and the initiation of tooth development was observed. However the recombination of second arch mesenchyme which expresses the *Lhx6* and *Lhx7* with the second branchial arch epithelium reduced the expression of both genes and led to failure of tooth development. Thus, *Lhx6* and *Lhx7* expressions need signals from the oral epithelium of the first branchial arch to be activated (Tucker and Sharpe, 2004, Nanci and Ten Cate, 2008).

The role of secreted fibroblast growth factor FGF8, in odontogenesis has been largely supported in the literature for its involvement in the early induction of *Lhx6 and Lhx7* (Tucker *et al.*, 1999). However, *Lhx6 and Lhx7* expression is not affected by *Bmp4* (Grigoriou *et al*., 1998). Tucker and Sharp have suggested that *Lhx6 and Lhx7* might be important in marking the position of tooth bud but not sole genes for tooth development (Tucker and Sharpe, 2004). The homeobox gene *Gsc* (goosecoid) is expressed after the *Lhx6 and Lhx7* but not in their region, and it is also induced by *Fgf8*. Tucker *et al.* (1999) have divided the jaw into tooth forming LHX-positive domain and a non-tooth forming GSC-positive domain (Tucker *et al.*, 1999). In the *Gsc* knocked out mouse, teeth develop normally, but the underling skeletal elements that support the teeth are absent (Yamada *et al.*, 1995).

The transcription factor gene *Islet*, which is a member of the LIM homeodomain family, is expressed in the distal oral epithelium and acts in a positive feedback loop with the BMP4 as they both induce the expression of each other. FGF8 expression in oral epithelium is positively controlled by the paired related homeobox gene *Pitx2*. *Fgf8* and *Bmp*4 are

reciprocally antagonist and the expression of *Pitx2* in the oral epithelium is initially controlled by their antagonistic signals as these genes act together in positive and negative feedback loops. In the *Pitx2* knockout mouse the expression of *Fgf8* is reduced but the expression of *Bmp4* is induced with tooth development being arrested at the invagination/bud phase. These genes work together in these loops for further definition gene expression patterns in the oral epithelium (Lu *et al.,* 1999; Tucker and Sharpe, 2004).

However this confirm that the odontogenic potential exist in the oral epithelium around day 11.5 especially that the recombination of oral epithelium with trunk ectomesenchymal cell enhanced the tooth development (Mina and Kollar, 1987; Lumsden, 1988). After day 11.5 the recombining experiments demonstrated that tooth developing potential transfer from the oral epithelium to the mesenchyme as it showed ability to induce tooth formation when combined with a non-oral epithelium (Nanci and Ten Cate, 2008). Expression of some genes in ectomesenchyme around day 11 in mice (approximately 7 weeks in humans), marks the initiation of the future tooth such as *Pax9* and *Activin-A*. *Pax9* gene was identified as one of the earliest mesenchymal genes that regulated positively by the *Fgf8* early in the odontogenesis process and negatively by bone morphogenetic proteins (B*mp*2 and B*mp*4). In mice experiments, *Fgf8*, *Bmp2* and *Bmp4* are expressed in non-overlapping sites to regulate the expression of the *Pax9* as it exists at sites where *Fgf8* is but not *Bmps* (Figure 1.3). However *Pax9* seems not to be essential for the initiation process as in knockout mice experiment, tooth buds form but arrest at this stage causing tooth agenesis (Peters *et al.,* 1998). This suggests that the FGF8, BMP2 and BMP4 might not be involved directly in tooth initiation. In the same period, other genes such as *pitx2*, *Dlx2*, *Msx1*, *Msx2,* and *Lef1*are expressed in oral epithelium (Figure 1.3). Up to now, the specific role of each gene in epithelium at this stage has not been identified yet and it's not clear if any of these genes is

involved directly in the regulation of the *Fgf8* or *Bmp* expression. Once the odontogenesis process has been initiated by ectomesenchyme, dental papillary cells maintain its development (Nanci and Ten Cate, 2008).

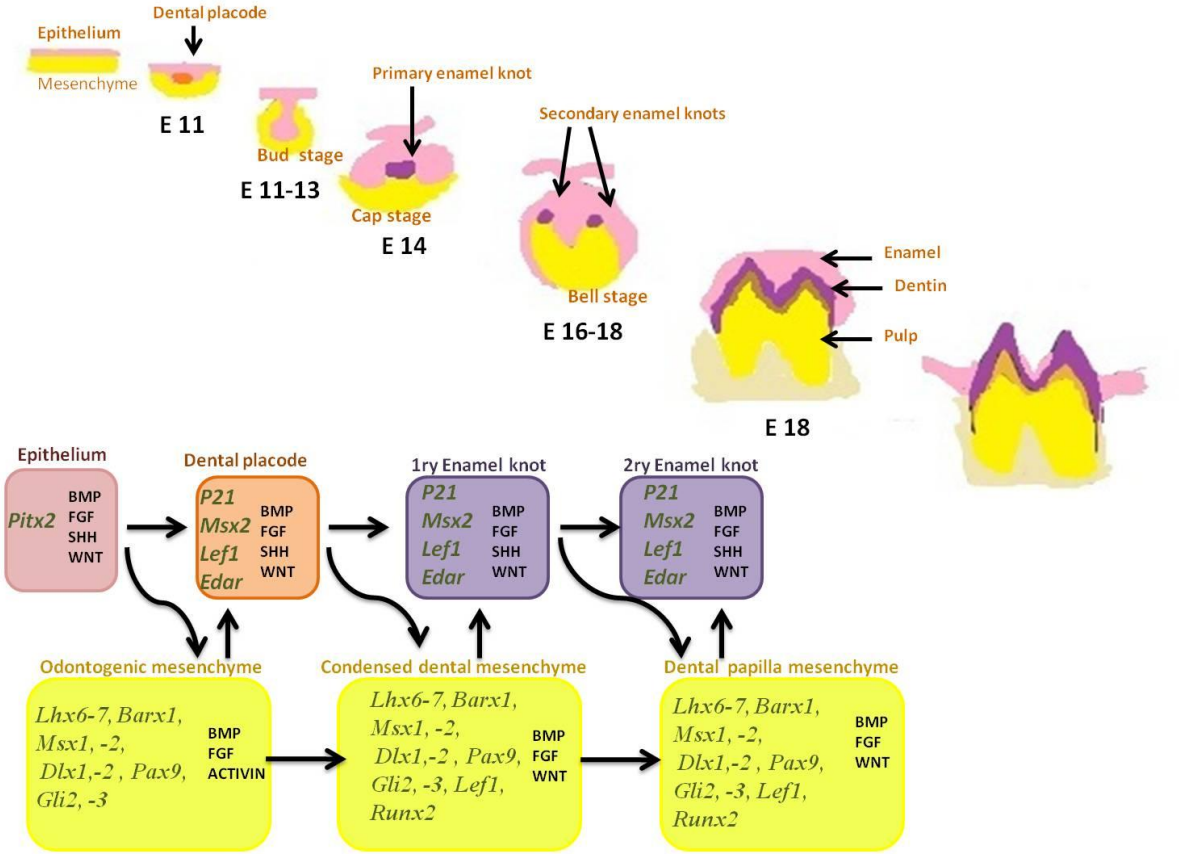


Figure 1.3: An overview of the reciprocal interactions between epithelium and mesenchyme during the early stage of tooth development in mice (E: embryonic development day. (Signalling **in bold and *transcription factors* in *italics*).**
Adapted from (Thesleff, 2006; Nanci and Ten Cate, 2008)

1.4.6 Tooth bud initiation

The initiation of the bud is followed by the induction of the enamel knot at the tip of epithelial bud, which expresses more than ten signal molecules for the transition of the epithelial bud to a cap (Figure 1.3). This transition marks the onset of tooth crown formation of different type of teeth. All the four signals families are engaged in the mediation of the

interactions. The enamel knot is regulated by Wnt and BMP signals. *Bmp4* induces *P21* expression which regulates the cell cycle arrest in the enamel knot. Wnts are essential for *Fgf4* expression in the knot (Thesleff, 2006).

Sonic hedgehog *Shh* has been reported as a regulator for the proliferation of dental epithelial cells to form the tooth bud (Hardcastle *et al*., 1998). However its direct effect seems to be on mesenchyme as it triggers the development of reciprocal signal acting back on the epithelium (Gritli-Linde *et al*., 2002). The inter action between the SHH and some of the WNT family genes sets up the boundary between the oral and the dental epithelia (Sarkar *et al.*, 2000). The adjacent tissue proliferation is regulated by the FGFs and their receptors in both epithelial and mesenchyme in reciprocal mode (Thesleff, 2006). During these stages of morphogenesis three transcription factors in the dental mesenchyme are expressed: *Msx1*, *Pax9* and *Runx2* and regulated by epithelial signals. As we mentioned in Section 1.4.3.6, *Pax9* mutants mice tooth bud arrested at bud phase (Peters *et al.*, 1998). Tooth development can also be arrested at bud phase by knocking out two genes in mice, *Msx1* and *Runx2*. The role of *Msx1* is not only in patterning the future incisor region, but also has a later role in the mesenchyme during the bud phase (Chen *et al.*, 1996). *Runx2-/-* knockout mice, arrested the tooth development at the bud stage only in the lower molars. The upper molars make it through to an abnormal cap stage with compensation effect of the *Runx3* which occurs only in the maxilla (Aberg *et al.*, 2004). The role of *Runx2* in tooth development extended from the bud stage to ameloblast in the dental mesenchyme during the maturation stage of enamel formation (D'Souza *et al.*, 1999) (Figure 1.3).

1.5 Mutations in humans

Mutations in transcription factors and signals genes have been identified in tooth agenesis patients with different phenotype and different mode of inheritance (Table 1.1): X linked (Suarez and Spence 1974; Han *et al.*, 2008), autosomal, dominant (Lammi *et al.*, 2003; Klein

et al., 2005; Kim *et al*., 2006) or recessive (Pirinen *et al.* 2001; Swinnen *et al*. 2008; Bloch-Zupan *et al*. 2011). Heterozygous loss of function of *MSX1* or *PAX9* leads to severe tooth agenesis (Vastardis *et al*., 1996; Stockton *et al*., 2000). The variation in phenotypes of both *MSX1* and *PAX9* mutations has been referred to the haplo-insufficiency that stops tooth development in an autosomal dominant manner (Kist *et al.*, 2005). Although the *Runx2* knocked out in mice resulted in tooth agenesis as in *Msx1* and *Pax9* mutant mice, a heterozygous loss of function in humans causes a condition called cleidocranial dysplasia which is a hereditary congenital disorder characterised by supernumerary teeth, delayed or non-eruption of teeth and skeletal defects (Hansen *et al.*, 2011, Mundlos *et al.*, 1997) which are also found in mice (Aberg *et al*., 2004). Since mice do not have second dentition but supernumerary teeth are found in this condition, these supernumerary teeth in cleidocranial dysplasia can be considered as a third dentition. Hence *Runx2* can be considered as a negative regulator of secondary teeth and positive regulator for the primary teeth (Thesleff, 2006).

Another transcription factor that has been reported with tooth agenesis is *AXIN2*. The mutation phenotype of this gene is associated with severe tooth agenesis that affects the secondary dentition and colorectal cancer (Lammi *et al*., 2004a; Callahan *et al*., 2009). However Mostowska *et.al* have identified three novel mutations in *AXIN2* in patients with non-syndromic isolated tooth agenesis (Mostowska *et al*., 2006a). *WNT10A* (Wingless-type MMTV integration site family member 10A) OMIM606268, mutations have been reported with different type of ectodermal dysplasia syndromes. Firstly Bohring *et al.*, proposed that heterozygous mutations in *WNT10A* might show isolated ectodermal developmental defects such as tooth agenesis in 53% of cases with a sex bias manifestation pattern of males being more affected than females (Bohring *et al.*, 2009). Then Kantaputra and Sripathomswat in 2011 reported a heterozygous mutation in an American family with isolated tooth agenesis

and microdontia (Kantaputra and Sripathomsawat, 2011). Interestingly *WNT10A* mutations have also been identified in isolated tooth agenesis with other pathogenic mutations in *MSX1*, *PAX9* and *AXIN2* (van den Boogaard *et al.* 2012). *EDA* gene, which is linked to X-linked hypoidrotic ectodermal dysplasia (HED), has been shown recently, with its receptor *EDAR* gene, to have link to families with isolated tooth agenesis (Tao *et al*., 2006; Han *et al*., 2008; Li *et al*., 2008; Bergendal *et al*., 2011).

More recently, affected children from a Turkish family of consanguineous marriage presenting with autosomal recessive oligodontia and extreme microdontia were reported to have a homozygous splice-site mutation in the *SMOC2* gene (SPARC related modular calcium binding 2 protein, OMIM 607223) at chromosome 6 (Bloch-Zupan *et al.*, 2011).

Table 1.1: **Known genes mutations in humans relevant to non-syndromic tooth agenesis**

Gene	Mutated amino acid	Phenotype (U=upper. L=lower)		Mode of inheritance	Reference
		Type	Unaffected teeth		
MSX1	Arg196Pro	Severe tooth agenesis	U central and lateral incisors, canines, L first premolars & second molars	Autosomal dominant	(Vastardis *et al.*, 1996)
MSX1	Ser105X	Severe tooth agenesis	Central incisors, L lateral incisors, canines & L first premolars	Autosomal dominant	(van den Boogaard *et al.*, 2000)
MSX1	Ser202X	Severe tooth agenesis	U central and lateral incisors	Autosomal dominant	(Jumlongras *et al*., 2001)
MSX1	Met61Lys	Severe tooth agenesis	Central incisors, L lateral incisors, canines and first molars	Autosomal dominant	(Lidral and Reising, 2002)
MSX1	Gln187stop	Severe tooth agenesis	Central incisors, canines, L first premolars, U first molars	Autosomal dominant	(De Muynck *et al.*, 2004)
MSX1	Gly22fsX168	Severe tooth agenesis	Upper incisors, L lateral incisors, canines, first molars	Autosomal dominant	(Kim *et al*., 2006)
MSX1	Ala149Val	Severe tooth agenesis	Teeth affected was not reported	Autosomal dominant	(Mostowska *et al.*, 2006)
MSX1	Ala219Thr	Severe tooth agenesis	U left central incisor, canines, & L first premolars	Autosomal recessive	(Chishti *et al.*, 2006)
MSX1	Ala 221glut	Severe tooth agenesis	First molars	Autosomal dominant	(Xuan *et al.*, 2008)
MSX1	Intronic deletion	Severe tooth agenesis	U central & lateral incisors, canines and first molars.	Sporadic	(Pawlowska *et al*., 2009)
MSX1	Lue224Pro	Selective tooth agenesis	Central and lateral incisors, canines, first premolars, first molars and second molars	Autosomal dominant	(Mostowska *et al*., 2012)
PAX9	Gly73fsX243	Severe tooth agenesis	U central incisors, canines, lateral incisors and first premolars	Autosomal dominant	(Stockton *et al.*, 2000)
PAX9	Lys114stop	Severe tooth agenesis	U central incisors and L first premolars	Autosomal dominant	(Nieminen *et al*., 2001)
PAX9	Val265fsX315	Severe tooth agenesis	U central incisors, canines, & L first premolars	Autosomal dominant	(Frazier-Bowers *et al*., 2002)
PAX9	Lys91glu	Severe tooth agenesis	U central incisors, lower canines & L first premolars	Autosomal dominant	(Das *et al.*, 2003)
PAX9	Leu21Pro	Severe tooth agenesis	U central incisors, lower canines & L first premolars	Autosomal dominant	(Das *et al.*, 2003)
PAX9	Arg59fsX177	Severe tooth agenesis	U central incisors, canines, lateral incisors & first premolars	Autosomal dominant	(Das *et al.*, 2003)
PAX9	Gly51Ser	Severe tooth agenesis	U central incisors, L canines, first	Sporadic	(Mostowska, Kobielak *et al*., 2003)

			premolars, and first and second molars		

Table .1.1. Continued

Gene	Mutated amino acid	Phenotype (U=upper. L=lower)		Mode of inheritance	Reference
		Type	Unaffected teeth		
PAX9	Arg26Tyr	Severe tooth agenesis	U central incisors, L canines & first premolars.	Autosomal dominant	(Lammi *et al*., 2003)
PAX9	Arg28Pro	Severe tooth agenesis	U central incisors, upper first premolars, L lateral incisors, L canines & L first and second premolars.	Autosomal dominant	(Jumlongras *et al.*, 2004)
PAX9	Met1Val	Severe tooth agenesis	U central incisors, L canines & L first premolars.	Autosomal dominant	(Klein *et al.*, 2005)
PAX9	Lys91Glu	Severe tooth agenesis	Central incisors, lateral incisors, canines & first premolars	Autosomal dominant	(Kapadia *et al.*, 2006)
PAX9	Arg59stop	Severe tooth agenesis	Central incisors, L lateral incisors, canines & L first and second premolars	Autosomal dominant	(Tallon-Walton *et al.*, 2007)
PAX9	Arg47Tyr	Severe tooth agenesis	L first and second premolars and L first molars	Autosomal dominant	(Zhao *et al.*, 2007)
PAX9	Gly6Arg	Selective tooth agenesis	U centrals, lateral incisors, canines, U first premolars, second premolars and first and second molars	Sporadic	(Wang *et al.*, 2009)
PAX9	Ser43lys	Severe tooth agenesis	Central incisors, L canines, U second premolars, L first and second premolars.	Autosomal dominant	(Wang, Wu *et al*. 2009)
PAX9	Ala168Gly	Severe tooth agenesis	Type of affected teeth was not reported	Autosomal dominant	(Boeira and Echeverrigaray, 2013)
AXIN2	Arg656stop	Severe tooth agenesis	U central incisors.	Autosomal dominant	(Lammi et al., 2004b)
AXIN2	Asn666stop	Severe tooth agenesis	U central incisors, upper canines and first and second molars	Sporadic	(Lammi et al., 2004a)
EDA	Arg65Gly	Severe tooth agenesis	First molars	X linked	(Tao *et al.*, 2006)
EDA	Thr338Met	Severe tooth agenesis	All types of teeth were affected by this mutation	X linked	(Han *et al*., 2008)
EDA	Asp316Gly	Severe tooth agenesis	All types of teeth were affected by this mutation	X linked	(Li *et al.*, 2008)
EDA	Thr338Met	Severe tooth agenesis	Molars	X linked	(Li *et al.*, 2008)
EDA	Met364Thr	Selective to severe tooth agenesis	Molars	X linked	(Rasool *et al.*, 2008)
EDA	Gln331His	Severe tooth agenesis	Canines	X linked	(Ayub *et al.*, 2010)
WNT10A	Phe228lle	Selective tooth agenesis	Central incisors, canines, U second premolars, first premolars and molars	Autosomal dominant	(Kantaputra and Sripathomsawat, 2011)
SMOC2	Splice-site intronic	Severe tooth agenesis	Upper central incisors, lateral incisors, canines and U and L first and second molars	Autosomal recessive	(Bloch-Zupan *et al*., 2011)

1.6 Dental anomalies associated with tooth agenesis

1.6.1 Reduction in tooth size (microdontia) and morphology

Reductions in the mesio-distal and bucco-lingual dimensions of tooth crowns (microdontia) (Figure 1.4) associated with different types and severity of tooth agenesis have been noted by numerous authors (Grahnen, 1956; Garn and Lewis, 1970; Brook, 1984; Lammi *et al.,* 2003; Brook *et al.,* 2009b). This observation has been reported more with third molar agenesis (Celikoglu *et al.,* 2011). The connection between agenesis of teeth and atypical morphology of remaining teeth has been investigated in several studies with third molars and other types of agenesis (Al-Emran, 1990; Salama and Abdel-Megid, 1994). Some authors found that reduction in tooth size and change in morphology, such as in the Carabelli cusp, was greater with third molar agenesis than in the other types of tooth agenesis (Garn and Lewis, 1970). The association between tooth agenesis and generalised microdontia was reported to be more common with severe types of tooth agenesis (oligodontia) (Buckley and Doran, 2001; Brook *et al*. 2009b; Bloch-Zupan *et al*., 2011), whereas in selective tooth agenesis, according to case reports, the reduction in tooth size and morphology was localised mainly in third molars and upper lateral incisors (Garib *et al*. 2009, 2010).

Baccetti (1998) showed significant reciprocal associations between agenesis of second premolars and reduced upper lateral incisors, the prevalence of second premolars agenesis in his sample (1000 subjects) was 5.8% and was 4.7% for the upper lateral incisors. The group with agenesis of second premolars (100 subjects) had a higher prevalence of small maxillary lateral incisors than the control group with p value < 0.005 (18/100 subjects) and conversely, the group with small maxillary laterals (100 subjects) had a higher prevalence of agenesis of second premolars than their control group (42/100 subjects) (Baccetti, 1998b). These observations support the hypothesis of the multifactorial model in which tooth agenesis and

microdontia lie at one end of the curve (Brook, 1984). Peg-shaped upper lateral incisors were found in 5.5% of the family members of the probands with tooth agenesis compared with a frequency of 1.7% in the general population (Grahnen, 1956). In Baccetti's study, agenesis of one lateral incisor was often accompanied by a small lateral incisor on the contra-lateral side (Baccetti, 1998b). Alvesalo and Portin (1969) suggested that agenesis and peg-shaping of upper lateral incisors might be due to an autosomal dominant gene with reduced penetrance (Alvesalo and Portin, 1969).

Figure 1.4: A clinical photograph of patient with association of tooth agenesis and microdontia.

1.6.2 Ectopic maxillary canines and other teeth

The prevalence in the general population with palatally displaced canines has been reported by many authors between 1 to 2% (Shah *et al.*, 1978; Grover and Lorton, 1985; Ericson and Kurol, 1986). Several attempts have been made to study the aetiology of mal-position canine, and some authors have suggested local factors are causative, such as lack of guidance in eruption due to diminutive lateral incisors (Becker *et al.*, 1981; Brin *et al.*, 1986; Oliver *et al.*, 1989).

However, a familial model has also been proposed (Zilberman *et al.* 1990; Peck *et al.*, 1994). Some investigators have suggested an association between selective tooth agenesis and ectopic position of teeth including canines (Figure 1.5), and maxillary permanent first molars (Svinhufvud *et al.*, 1988; Bjerklin *et al.*, 1992; Pirinen *et al.*, 1996).

Figure 1.5: An orthopantomogram radiograph shows an association between tooth agenesis of upper lateral and lower incisors and an impaction of upper right canine (white arrow)

1.6.3 Infra-occlusion of primary molar (s)

The tooth malposition shown in Figure 1.6 has been noted with agenesis of the second premolar. Baccetti and his co-workers pointed out the association between infra-occlusion of primary molars and agenesis of premolars with a prevalence of about 18% 22% (Baccetti *et al.,* 1998b) whereas a prevalence of infra-occlusion of 10% has been reported in another study (Bjerklin *et al.*, 1992).

Figure 1.6: A clinical photograph of an infra-occluded lower molar (black arrow).

1.6.4 Delayed formation and eruption of teeth

Garn and his co-workers noticed that agenesis of the lower third molar (s) or third molar (s) together with some other teeth, were associated with delayed formation and eruption of

premolars and molars (Garn *et al.,* 1961). Rune and Sarnäs (1974) confirmed this observation by reporting a mean of 1.8 years' delay for boys and 2.0 years for girls in children with missing up to seven teeth including the third molars (Rune and Sarnas, 1974). The delay in development or eruption of teeth has been reported with agenesis of other teeth such as the lower second premolars (Uner *et al.,* 1994).

1.6.5 Short roots of teeth

Generalised root shortening is not common, but there are some studies on localised root shortening, especially short-rooted maxillary central incisors with a prevalence of about 2.4 to 2.7% in Caucasians and as high as 10% in Mongolian populations (Apajalahti *et al.,* 1999). The association between tooth agenesis and root shortening of different tooth groups has been reported by Apajalahti *et al.,* (1999), and the most affected teeth were reported, in order, as: maxillary central incisors, maxillary premolars, mandibular premolars, maxillary canines, mandibular central incisors, maxillary lateral incisors, and first molars. This phenomenon occurred in 46% of their sample (Apajalahti *et al.,* 1999). The association between tooth agenesis and root shortening has also been mentioned in earlier studies (Lind 1972; Edwards and Roberts, 1990).

1.6.6 Rotation of non-adjacent teeth

Baccetti made an interesting observation of significantly high prevalence of tooth rotation in association with agenesis of nonadjacent teeth (Figure 1.7). He concluded that rotation of premolars is significantly associated with absence of upper lateral incisors. In unilateral agenesis of upper lateral incisors the rotation of the lateral incisor on the contra lateral side of the same arch has been reported. This observation holds true for unilateral agenesis of premolars and rotation of premolars on the other side of the same arch (Baccetti, 1998a).

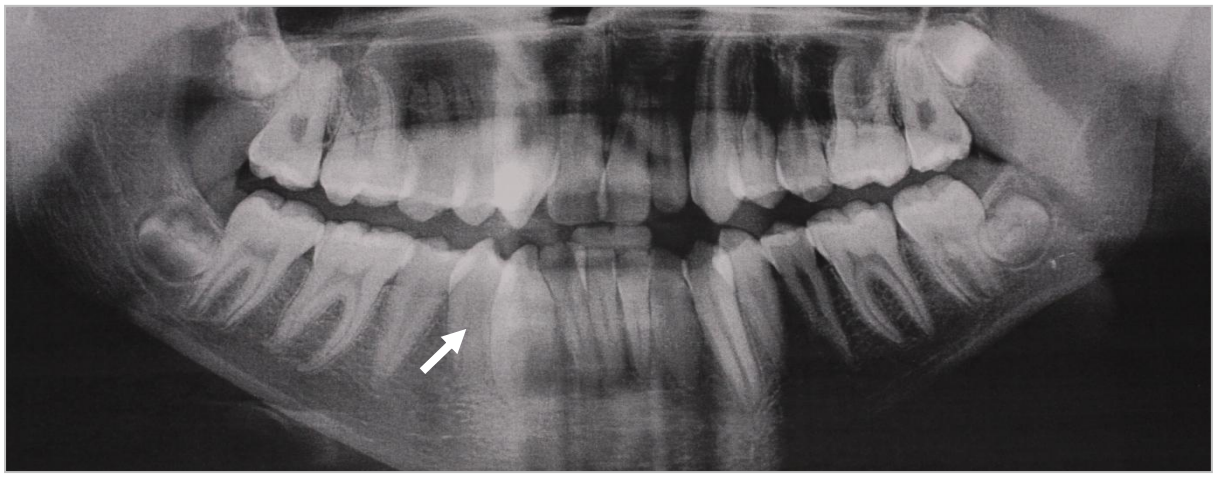


Figure 1.7: An orthopantomogram radiograph shows an association between tooth agenesis of upper right lateral incisor and rotation of the lower first premolar and the canine on the left side (white arrow).

1.6.7 Taurodontism

Taurodontism is a developmental anomaly that involves molar teeth in which the bifurcation or trifurcation of roots is close to the apex, with an abnormally large and long pulp chamber (Figure 1.8). Taurodontism has been reported by many authors as associated with tooth agenesis (Stenvik *et al.*, 1972; Lai and Seow, 1989; Schalk-van der Weide, *et al.* 1993; Seow 1993; Gomes *et al.*, 2012). However Calvano Kuchler *et al.*, (2008) in their study disagreed with previous studies and concluded that selective tooth agenesis was not associated with selective taurodontism in their data (Calvano Kuchler *et al.*, 2008).

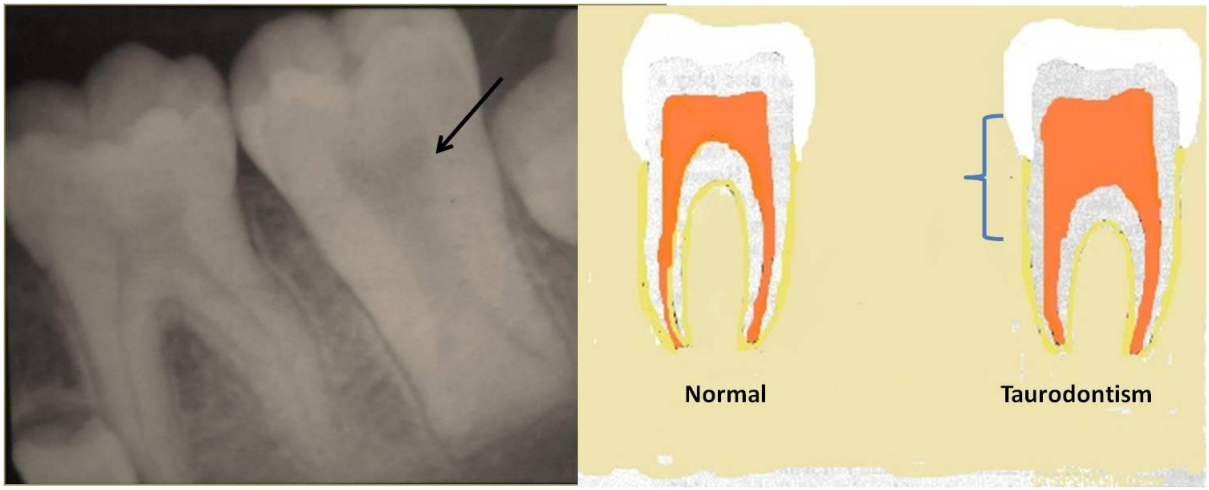


Figure 1.8: A comparison between the pulp chamber of a normal tooth and a tooth with taurodontism (black arrow).

1.6.8 Enamel hypoplasia, hypocalcification

Enamel hypoplasia and hypocalcification in association with tooth agenesis (Figure 1.9) has been pointed out by some authors. Ahmad *et al.,* (1998) investigated a family with a recessive inheritance model of tooth agenesis, and reported that the affected individuals had associated dental anomalies such as enamel hypoplasia, hypocalcification, and dentinogenesis (Ahmad *et al.,* 1998). Baccetti has included enamel hypoplasia as a dental anomaly associated with tooth agenesis (Baccetti, 1998b). Recently a Turkish family with severe tooth agenesis linked to a recessive mutation in *SMOC2* has been reported with tiny globular malformed crowns (Bloch-Zupan *et al.,* 2011).

Figure 1.9: A clinical photograph of a patient with tooth agenesis and enamel hypoplasia.

1.7 Tooth agenesis associated with isolated clefting and syndromes

150 or more syndromes have tooth agenesis as a part of their phenotype, most of them are listed in the British Baraitser-Winter Dysmorphology Database (BWDD).

1.7.1 Isolated cleft lip and/or palate

The incidence of tooth agenesis in association with different forms of cleft lip, cleft palate and both has been reported in numerous studies (Ranta, 1986; Shapira *et al.,* 2000b; Lourenco Ribeiro *et al.,* 2003). When compared with the general population, people with a cleft lip and palate have been found to have a higher prevalence of tooth agenesis that ranges from 10% to 77% especially in the area of cleft defect, this data fit with what was

hypothesised in the anatomical model (Svinhufvud *et al.*, 1988). The upper permanent lateral incisors were the most affected teeth in the cleft area, followed by the upper permanent second premolars and lower permanent second premolars, and the teeth that were most often missing on the non-cleft side (Ranta and Tulensalo, 1988) were the upper second premolars. Tooth agenesis was found more on the left side as it was the most affected side in clefting (Shapira *et al.* 1999, 2000b). Interestingly Lekkas *et al.*, (2000) found when they examined un-operated adult patients with cleft that there were no permanent teeth missing in the maxillary arch outside the cleft area (distal to the canines). They propose that the surgical procedure to close palatal clefts might be the reason for high prevalence of tooth agenesis as it disrupts the formation of the developing tooth buds in the upper arch (Lekkas *et al.*, 2000). Slayton *et al.*, (2003) noted that the data showing individuals with cleft palate have a greater prevalence of tooth agenesis in areas outside the cleft, compared with control individuals, is less conclusive (Slayton *et al.*, 2003).

1.7.2 Ectodermal dysplasias

Ectodermal dysplasias (EDs) as defined by Freire-Maia (1971), are congenital disorders characterised by alterations in two or more ectodermal tissues, at least one of these involving alterations in hair, teeth, nails or sweat glands (Freire-Maia, 1971). EDs are a large group of heterogeneous heritable conditions that involve congenital defects of one or more ectodermal structures and their appendages: hair (hypotrichosis, partial or total alopecia), nails (dystrophic, abnormally keratinised), teeth (enamel defect or tooth agenesis), sweat glands (hypoplastic or aplastic) and other structures derived from embryonic ectoderm such as mammary glands, the thyroid gland, thymus, anterior pituitary, adrenal medulla, central nervous system, external ear, melanocytes, cornea conjunctiva, lacrimal gland, and lacrimal duct (Itin, 2009).

In 1971 only eight types of ectodermal dysplasia were reported, nowadays this number has considerably increased to about 200 and the causative genes of 30 have been identified. This reflects the great progress in medical genetics over the past decades (Freire-Maia *et al.*, 2001). Two different classifications have been proposed so far, Pinheiro and Freire-Maia (1994) used clinical features for their classification. Group A comprises conditions with signs affecting at least two of the classical structures: hair, teeth, nails and sweat glands (pure ED). Group B comprises conditions involving one of the four classical signs in association with another ectodermal sign, with 11 clinical subgroups (Pinheiro and Freire-Maia, 1994). The second classification, by Priolo, combines molecular genetics and clinical aspects. (Priolo and Lagana, 2001; Irvine, 2009). EDs are inherited as X-linked, autosomal dominant or recessive (Salinas *et al.* 2009).

1.7.2.1 Hypohidrotic ectodermal dysplasia

Hypohidrotic ectodermal dysplasia (EDA or HED) OMIM 305100 is the most common type of EDs, usually inherited as an X linked trait as a result of mutation in the *EDA1* gene located on chromosome Xq12-13.1 (Kere *et al.,* 1996; Monreal *et al.,* 1998; Zhang *et al.,* 2009; Wu *et al.,* 2012). Mutation in the EDA receptor encoding gene *EDAR*, located on chromosome 2q11-q13, and in the EDAR-Associated Death Domain encoding gene *EDARADD* located on chromosome 1q42-q43, (see Section 1.4.3.5) have been reported with an autosomal recessive and dominant HED (Headon *et al.,* 2001; Bal *et al.,* 2007; Naqvi *et al.,* 2011; Haghighi *et al.,* 2012). In the X-linked type, both males and females have tooth agenesis but affected males show the classical phenotype of EDs and female EDA mutation carriers have milder variable phenotypic expression. Unlike the X-linked EDA, the autosomal recessive/dominant heterozygous carriers show no features of EDA disorder (Lu and Schaffer, 2008).

1.7.2.2 Ankyloblepharon-ectodermal defects-clefting

Hay-Well syndrome, also known as Ankyloblepharon-ectodermal defects-clefting syndrome (AEC) OMIM 106260, is a rare autosomal dominant disorder, characterised by dry and fragile skin, ankyloblepharon, tooth agenesis, orofacial clefting, alopecia or spare hair and nail dystrophy. Heterozygous mutations in the transcription factor *P63* gene are the major cause of AEC (McGrath *et al.,* 2001). Clements *et al.* reported that the mutation in the *P63* gene was associated with reduced expression of keratin genes in skin samples of AEC patients (Clements *et al.*, 2012).

1.7.2.3 Cleft lip/palate-ectodermal dysplasia syndrome

Cleft lip/palate-ectodermal dysplasia syndrome (CLPED1) OMIM 225060 is transmitted as an autosomal recessive trait, characterised by scanty eyebrows and eyelashes, spare scalp hair, cleft lip/palate nail dysplasia and upper lateral tooth agenesis. The causative gene in this syndrome is Poliovirus receptor related-1 *PVRL1* gene located at chromosome 11 (11q23-q24) (Suzuki *et al.,* 2000). *PVRL1* encodes a cell adhesion molecule called nectin-1, which has been found highly expressed in the medial edge epithelium of the mouse embryo developing palate (Suzuki *et al.*, 2000).

1.7.2.4 Odonto-onycho-dermal dysplasia (OODD)

Odonto-onycho-dermal dysplasia OMIM 257980 is a rare autosomal recessive syndrome characterised by severe tooth agenesis, nail dystrophy, smooth tongue, dry skin, keratoderma and hyperhydrosis of palms and soles. Loss of function and missense mutations in the *WNT10A* gene located on chromosome 2q35 have been reported with this type of ED (Adaimy *et al.*, 2007; Nawaz *et al.*, 2009).

1.7.2.5 Ectrodactyly-ectodermal dysplasia-clefting

Ectrodactyly-ectodermal dysplasia-clefting (EEC) OMIM 604292 is an autosomal dominant disorder characterised by deformities of hand or leg and fingers (ectrodactyly), signs of

ectodermal dysplasia (tooth agenesis affecting only the permanent dentition) and cleft lip and/or palate (Buss *et al.*, 1995). A region in chromosome 17 (7q11.2-q21.3) has been linked to the EEC. Several different mutations of transcription factor *P63* gene have been identified in EEC families (Celli *et al.*, 1999) and recently another missense mutation in the *P63* gene in a Chinese family with EEC has been reported (Wei *et al.*, 2012)..

1.7.2.6 Oral-facial-digital syndrome type 1

Oral-facial-digital syndrome type 1 (OFD1) OMIM 311200 was first reported by Papillon-League and Psaume in 1954. It is an X-linked dominant condition characterised by malformations of the face, oral cavity, and digits. Tooth agenesis of the lower lateral incisors in this disorder is associated with a fibrous band in the same region (Ferrante *et al.*, 2001).

1.7.2.7 Witkop Tooth-Nail Syndrome

Tooth-nails syndrome (TNS) OMIM189500 was first described by Witkop in 1965. It is a rare autosomal dominant disorder characterised by tooth agenesis and dysplastic nails. The phenotype of this syndrome includes tooth agenesis, malformation of both dentitions and dysplasia of the nails, mainly the toenails. A nonsense mutation in *MSX1* gene has been linked to this type of ED in three generations of a family (Jumlongars *et al.*, 2001).

1.7.3 Incontinentia pigmenti (IP, Bloch-Sulzberger syndrome)

Incontinentia pigmenti (IP, Bloch-Sulzberger syndrome) OMIM 308300 is an X-linked dominant disorder characterised by abnormalities in the skin with vesicular, verrucous and pigmented macular lesions, in hair, nails, central nervous system and teeth (severe tooth agenesis in both dentitions, microdontia, macrodontia, and taurodontism). This disorder has been linked to mutation in the *IKK-gamma* (NEMO) gene located in Xq28 (Smahi *et al.*, 2000).

1.7.4 Steatocystoma multiplex

Steatocystoma multiplex (SM) OMIM184500, is a rare condition that occurs at puberty with multiple skin-coloured nodules, caused by mutation in the keratin 17 gene (*KRT17*) (Smith *et al.*, 1997). A 2009 case report describes a 17-year old man presenting with multiple papules of the axillae, groin, neck and upper chest and tooth agenesis of seven teeth with a heterozygous missense mutation in *KRT17* and a role for the *KRT17* in tooth agenesis was suggested (Gass *et al.*, 2009).

1.7.5 Axenfeld-Rieger syndrome

Axenfeld-Rieger syndrome (ARS) OMIM180500 is a rare heterogeneous autosomal dominant disorder, characterised by malformations of the anterior chamber of the eye, umbilical, and craniofacial regions with mainly maxillary hypoplasia and dental anomalies such as tooth agenesis, mainly of the upper deciduous and permanent incisors and second premolars. Microdontia, taurodontism and short roots have also been reported with this condition. ARS has been linked to mutations in *FOXC1* located at chromosome 6p25, and *PITX2* at chromosome 4q25. A third locus for ARS was mapped to chromosome 13q14 but the relevant gene has not yet been identified (Dressler *et al.*, 2010).

1.7.6 Van der Woude syndrome

Van der Woude Syndrome (VWS) OMIM119300 is an autosomal dominant disorder associated with cleft lip and/or palate, pits of the lower lip and tooth agenesis. The interferon regulatory factor-6 (*IRF6*) gene located at chromosome 1q32.2, has been linked to VWS (Desmyter *et al.*, 2010).

1.7.7 MSX1 mutation

A mutation in exone 1 of *MSX1* in chromosome 4 has been reported in a large Dutch family with tooth agenesis and various combinations of cleft lip and palate (van den Boogaard *et al.*, 2000).

1.7.8 Down syndrome

Down syndrome OMIM190685, is the most common chromosomal abnormality in man, and is caused by trisomy of all or a critical part of chromosome 21 (21q22.3). Down syndrome is characterised by a group of phenotypic features including dysmorphic features and mental retardation. Congenital malformations of the heart occur in about 40% of the patients. Tooth agenesis has been reported in about 23-47% of cases. The most affected teeth are reported as the upper lateral incisors, lower incisors, second premolars and third molars (Shapira *et al.*, 2000a)

1.7.9 Holoprosencephaly

Holoprosencephaly (HPE) OMIM236100 is a rare heterogeneous malformation sequence in which the basic feature is impaired midline cleavage of the embryonic forebrain. It is characterised by facial dysmorphism including cyclopia, hypertolirism, single nostril, cleft lip and tooth agenesis. Generally, 3 degrees of severity have been described, defined by the extent of brain malformation. Several loci for holoprosencephaly have been mapped to certain chromosomal sites and the causative genes in some of these loci have been identified, such as *SHH, TGIF, and SIX3* genes (Wallis and Muenke, 2000).

1.7.10 Diastrophic dysplasia

Diastrophic dysplasia (DTD) OMIM222600 is an autosomal recessive disorder. Its main features are short-limbed short stature, generalised joint dysplasia and spinal deformities. The oral phenotype is tooth agenesis in the permanent dentition mainly: a) the lower second

premolar, upper lateral incisor and upper second premolar (in 30% of cases), or b) microdontia, cleft palate or submucous cleft palate (in a further 30%). DTD has been linked to the sulphate transporter gene *DTDST* in chromosome 5 (Dwyer *et al.*, 2010).

1.8 The Human Genetics

1.8.1 The complex genome

The genome holds 3.2 billion nucleotides, encoding 22,000 genes and various genomic components that participate in regulation of gene expression, such as histone modification, epigenetics, splice variants, micro RNAs and long noncoding RNAs. Each gene contains regulatory region exons (the protein coding regions), and introns (the intervening segments between exons). The introns occupy approximately 5% of the genome and the protein-coding exons (exome) occupy only approximately 1% (Lander, 2011). The function of about 98% of the genome up to 2003 was unknown (Dunham *et al*., 2012), and was considered as non functional. However, the "DNA junk" concept does not exist anymore since the encyclopaedia of DNA elements (ENCODE) project lunched by the National Human Genome Research Institute (NHGRI) revealed that, the vast majority (80.4 %) of the genome was found to have various biochemical activities such as RNA transcription, transcription factor binding, chromatin structure, and histone modification in at least one cell type (Dunham *et al*., 2012; Khatun *et al.*, 2013). The ENCODE group concluded that the majority of the human genome is involved in different biochemical functions (Doolittle, 2013).

1.8.2 Diversity of the human genome

Humans are genetically diverse; no two humans have identical genomes with the exception of identical twins. However, the differences were only in approximately 0.1% of our genomes. The single nucleotide polymorphism (SNP) database (dbSNP, Build 132) lists more than 37 million variants amongst humans. Each genome includes at least 4 million genetic variants

and many of them are personal (Gamazon *et al*., 2010). Most genetic variants in the genome are SNPs, but structural variations (SVs) which contain deletions, insertions, duplications, and arrangements of large segments of DNA affect more nucleotides in the genome. Interestingly each genome has approximately 50 to 100 variants that have been associated with inherited disorders and about 30 *de novo* variants (Marian, 2012b).

1.8.3 Mutations

A mutation is defined as a structural change in genomic DNA which can be transmitted from a cell to its daughter cells. Described below is the structural classification of mutations.

1.8.3.1 Point mutations/ substitution

A substitution mutation contains a change to a single nucleotide. Depending upon the effect on the protein product, substitution mutations can be subdivided into silent, missense, nonsense, regulatory and RNA processing mutations.

Silent mutations do not alter the amino acid being encoded.

Missense mutations result in the substitution of one amino acid for another. They account for about 45-50% of all known pathogenic human mutations (Young, 2005).

Nonsense mutations create a new stop codon (UAA, UAG, or UGA) which results in premature termination of translation. Thus this mutation is also known as premature termination mutation; it accounts for 11% of all known pathogenic human mutations.

1.8.3.2 Deletions and insertions

This type of mutation is sub-divided based on size into:

Small deletions and insertions

These involve loss (deletion) or gain (insertion) of a few nucleotides. If the number of nucleotides deleted or inserted in an exon is not a multiple of three, then the sequence of codons, known as the reading frame, will be disrupted. This is referred to as a frame-shift and

it usually results in a shortened (truncated) protein product. Small deletions and insertions account for about 22% of all mutations (Young, 2005).

Large deletions and insertions

These range in size from 20bp to 10Mb. More than that, the mutation will be visible under a light microscope and is classified as a chromosome abnormality. These mutations account for approximately 5-6% of all known pathogenic human mutations (Young, 2005).

1.8.4 Single-gene (Mendelian) and polygenic disorders and type of inheritance

Polygenic inheritance results from the combined effects of several genes, with small contributions from each and possibly an influence from environmental factors, whereas single-gene (Mendelian) disorders are caused by a mutation in a single gene (Strachan and Read, 2010).

1.8.4.1 Autosomal dominant inheritance

In this mode of inheritance the disorder is caused by an error in a single copy of a gene located in one of the autosomal genes. Thus the affected individual will carry one wild type allele (normal) and one abnormal or mutant allele. Most individuals affected with an autosomal dominant disorder are heterozygotes. Rarely, homozygosity can occur, mostly because of mating between two heterozygotes, and results in a severe phenotype (Strachan and Read, 2010). Variability in the expression of the gene might occur between individuals in the same family. In rare cases, a disorder with autosomal dominant inheritance may appear to skip a generation and shows absolutely no manifestations (non-penetrance) (Young, 2005), or variable penetrance, for example, in arrhythmogenic right ventricular cardiomyopathy (ARVC) (Murray, 2012) and Ichthyosis vulgaris (Thyssen *et al*., 2013).

1.8.4.2 Autosomal recessive inheritance

Disorders which show autosomal recessive inheritance occur only in the homozygous and compound heterozygote states, in which the affected individual has two mutant alleles inherited from each heterozygous parent (who is a carrier) (Strachan and Read, 2010). An example of this is cystic fibrosis (CF) OMIM 219700.

1.8.4.3 Semi-dominant inheritance

Semi-dominant inheritance refers to the condition where autosomal recessive disorders show an autosomal dominant pedigree. This situation occurs in consanguineous families with a high incidence of carriers within the extended family (Young, 2005). An example of this is Alzheimer disease (AD) OMIM 104300

1.8.4.4 X linked inheritance

X linked disorders dominant and recessive are caused by a mutation in a gene located on the X chromosome and the manifestations of the disease usually occur in males, while females appear to be normal as they have two X chromosomes and the normal allele compensates for the effects of the mutant allele (Strachan and Read, 2010). An example of this is haemophilia.

1.9 Positional cloning of disease genes

Cloning of disease genes can be achieved through two processes, functional and positional cloning. Functional cloning, as the name implies identifies the disease gene based on knowledge of gene products (encoded protein), such as the protein amino acid sequences and the biochemical defect of the disease. If this can be done, gene-specific oligonucleotides or specific antibodies will be used to identify the gene, this method has been successfully used to identify the responsible gene of several diseases such as the haemophilia A gene (Pio *et al.*, 2009). As the biochemical basis of the tooth development is still elusive, the use of functional

cloning in familial tooth agenesis is not possible. Thus this thesis will focus on positional cloning as a successful approach in identifying causative genes in Mendelian disorders.

1.9.1 Positional cloning

Positional cloning is based on the cloning or isolation of a gene through knowledge of its location in the human genome. Revolutions in modern technology and information gained from the Human Genome Project have united to make positional cloning a tremendously powerful method for identifying causative genes in single gene disorders. The advance development of exome sequencing has made many other approaches obsolete such as DNA microarrays which are used for linkage analysis. The positional cloning process involves two main steps: first defining the approximate gene location, then identifying the candidate genes (Strachan and Read, 2010).

1.9.2 Next generation sequencing (high throughput sequencing)

1.9.2.1 Whole exome sequencing

The current improvements in high-throughput sequence capture methods and next generation sequencing (NGS) technologies have made exome sequencing feasible and cost effective (Table 1.2). This method has become a powerful new approach for identifying the disease genes in monogenic Mendelian and complex disorders. The steps of the whole exome sequencing are illustrated in Figure 1.10. The strength of this method comes from the broad discovery of protein-coding variants throughout the genome, which means that even small families and isolated affected individuals, can now be used to discover the disease mutations by using whole exome sequencing.

Table 1.2: A comparison of Next Generation Sequencing instruments (Majewski *et al*., 2011).

Instrument	Current max output	Cost per 1 million Bases (in US$)
Roche GS-FLX	1.5M reads/run (1 day) 500bases/read	$10
Life Technologies Ion Torrent	100k reads/run (2 h) 100bases/read	$1
Life Technologies SOLiD5500XL	4G reads/run (7days) 75 bases/run	$0.13
Illumina HiSeq 2000	6G reads/run (8 days) 100bases/read	$0.05-$0.15
Pacific Biosciences	100k reads/run (2 h) >1000bases/run	$2

The basic steps required for exome sequencing are shown in Figure 1.10. Exome sequencing has identified about 20,000 single nucleotide variants (SNVs); most of these variants (95%) have been identified as polymorphisms in the human population. The strategies used with exome sequencing to identify the causal alleles are dependent on many factors such as the mode of inheritance of the trait and pedigree structure, and sample size. Identifying the causal alleles in exome sequencing relies on discrete filtering of the data to reduce the number of candidate genes to a minimum number of the most likely candidates, and comparisons with exome sequences and variants that are found in a small number of unrelated or closely related affected individuals to identify rare or novel variants in the same gene shared between affected individuals. The novelty of the variants is assessed by filtering the variants against a set of polymorphisms available in public databases such as dbSNP and 1000 Genome project, or those found in a set of unaffected individuals. About 2% of the SNVs identified in an individual by exome sequencing are novel (Bamshad *et al.,* 2011; Singleton, 2011). After discrete filtering, the next step is to stratify candidate genes by their functional class, focusing on the frameshifts, stop codons and disruption of canonical splice site variants. Candidate variants can also be stratified by existing biological or functional knowledge about a gene.

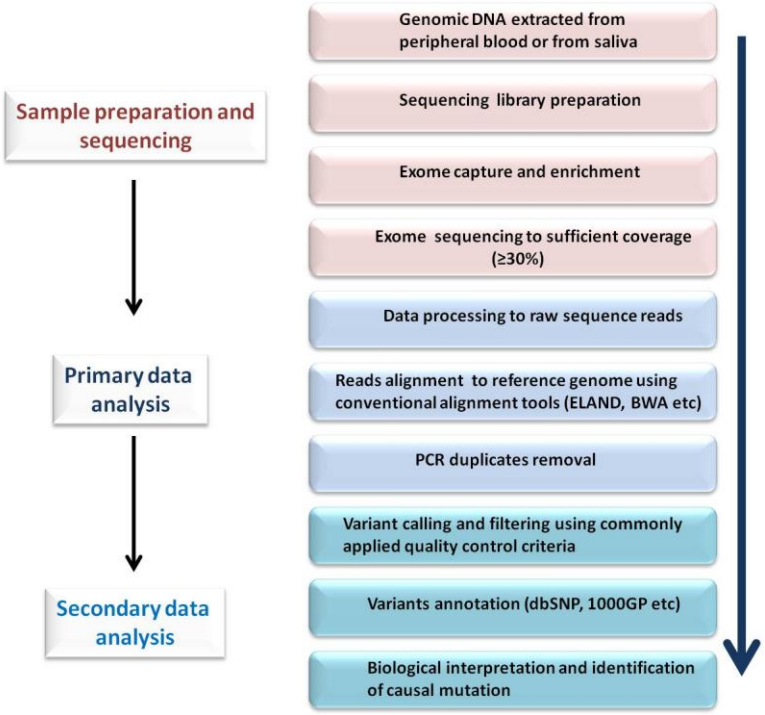


Figure 1.10: Work flow of exome sequencing from sample preparation to the identification of the causal mutation (Ku *et al*., 2012).

1.9.2.2 Whole genome sequence

Whole genome sequencing is a genetic laboratory process used to assess the complete DNA sequence of the genome at one time. It is a very powerful technique that has been used recently to investigate complex disorders. However, several complex diseases have been successfully investigated by homozygosity mapping and whole genome sequencing, for example, autism, mental retardation, epilepsy, schizophrenia (Stankiewicz and Lupski, 2010) and cancer (Majewski *et al.,* 2011). The limitation in the use of this method is its cost, which is very high compared to other approaches such as exome sequencing. The main advantage of the whole genome sequencing is its ability to study genetic variants that are not present in the exomes. However there are some authors who stress the advantages of whole exome sequencing in comparison with whole genome sequencing, based on the fact that most known causal variants for Mendelian disorders are found in exons. In addition, although whole genome sequencing has been performed in some studies for Mendelian disorders, these analyses still focus on the variants in exones (Ku *et al*., 2011).

1.10 Summary of literature review

A considerable amount of literature has been published on non syndromic tooth agenesis. The prevalence of this condition which is a developmental failure of either deciduous or permanent teeth, in general populations, was range from 1.6 to 9.6% for permanent dentition (Nieminen, 2009). In deciduous dentition the prevalence was less than 1% (Arte *et al.*, 2001).

Teeth develop through successive bud, cap, and bell stage any intervention or defect in these stages will cause tooth agenesis. The molecular basis of tooth development has been studied in mice and human models. These studies have revealed that tooth development is under strict genetic control; almost three hundred genes have been identified in odontogenesis processes. These include signals such as BMP, FGF, Shh and Wnt, and transcriptions factors in the MSX, DLX, LHX families and PAX9 (Thesleff, 2006).

Up to now the majority of the mutations that have been linked to non syndromic tooth agenesis in human were in *MSX1* and *PAX9* (Table 1.1), but there were another studies verified that mutations and variants in *EDA*, *AXIN2*, *SMOC2,* and *WNT10* genes also associated with tooth agenesis (Han *et al.,* 2008, Nieminen, 2009; Bloch-Zupan *et al.,* 2011; Kantaputra and Sripathomsawat, 2011).

Non syndromic tooth agenesis has been reported with association of other dental anomalies such as microdontia, ectopic maxillary canine, infra-occlusion of primary molar, delayed formation and eruption of teeth and taurodontism (Baccetti, 1998).

Different phenotypes of tooth agenesis with autosomal dominant, recessive X link and sporadic mode of inheritance have been reported in the literature (Lidral and Reising, 2002, Das et al., 2003, Lammi et al., 2004b, Chishti et al., 2006, Rasool et al., 2008). However, only two studies were found to be in families of consanguineous marriage.

Recently, there has been a rapid development on gene mapping techniques, especially with exom sequencing. Hence, the main aim of this study was to use the conventional and the novel gene mapping techniques to locate the autosomal recessive gene of tooth agenesis.

1.11 Aims and hypothesis

The general aims of this study were:

- to investigate the phenotypic and genotypic features of non syndromic familial tooth agenesis in Saudi and Pakistani consanguineous families.
- to explore the type and segregation of the causative mutations of tooth agenesis in theses unique.
- to study the broad spectrum of tooth agenesis phenotype in these families.
- to carry out an initial study of tooth crown dimensions in relation to tooth agenesis in one large Saudi family.

This is the first study to investigate the molecular basis of tooth agenesis in Saudi families.

The aims and objectives for each results chapter of this thesis are described below:

1.11.1 Chapter 2: The clinical characterisation of families with tooth agenesis

The aim of this chapter was to investigate the segregation and phenotype of tooth agenesis in 16 families presenting different phenotypes of the condition by:

- characterising the mode of inheritance and variability in phenotypes among families with different types of tooth agenesis by clinical and radiographic assessments.
- classifying the tooth agenesis based on the phenotype.
- reporting and describing all dental anomalies associated with tooth agenesis in the family members.

1.11.2 Chapter 3: The genetic investigations to identify the diseased mutations in families with tooth agenesis

This chapter aimed to identify the causal mutations in eight consanguineous families by

- identifying a large homozygous region(s) shared by all the affected members in one family, or shared by some of the affected members in different families.
- discovering a rare variant(s) segregating with the disease (tooth agenesis) in the affected members.

1.11.3 Chapter 4: Tooth crown dimensions in Family I (initial study)

This initial study aimed to determine whether there are any evident patterns or trends in the dental crown sizes of the remaining teeth in *Family I*, which has an autosomal recessive mild to moderate tooth agenesis, with lower second premolars as the most affected teeth (see Section 2.3.3.1).

- It was hypothesised that patients with tooth agenesis may have a reduction in their crown dimensions (microdontia) as these two dental anomalies may have a similar aetiology. To test this hypothesis manual mesio-distal and bucco-lingual measurements on study model dental casts of patients with tooth agenesis and their close relatives from one large family with non-syndromic selective tooth agenesis were recorded and compared with a control group from the same region.

CHAPTER 2: THE CLINICAL CHARACTERISATION OF FAMILIES WITH TOOTH AGENESIS

Tooth agenesis is often used as a collective term for congenitally missing teeth. Agenesis of one or a few teeth is the most common clinical condition in dental clinics. This type of dental anomaly may cause aesthetic and functional impairment, whereas relatively rare severe forms definitely need multidisciplinary dental care. Phenotype and genotype characterisation of familial tooth agenesis should help dentists to deliver effective counselling for patients with tooth agenesis.

Although most reported tooth agenesis cases are autosomal dominant in inheritance (Kim *et al.,* 2006), an autosomal recessive pattern of inheritance has been reported several times in highly consanguineous populations (Chishti *et al.,* 2006; Bloch-Zupan *et al.,* 2011). In Saudi and Pakistani populations, consanguineous marriage is common and autosomal recessive inheritance is expected. However, in the Saudi population with its relatively homogenous genetic pool, no study has been carried out to locate the gene in a family with tooth agenesis.

2.1 Methods

2.1.1 Ethical approval

The experiments were approved by the ethics committee of the National Guard (Health Affairs) (Appendix A, Figure1) and the Queen Mary University of London (QMUL) research ethics committee (Appendix A, Figure2). Written and oral consent for gene screening was obtained from all participants. The use of the DNA samples in this study was complied with United Kingdom and Saudi Arabia regulations. The project was funded by the Saudi Ministry of Higher Education.

2.1.2 Patients database

Patients were recruited from Saudi dental clinics that provide dental care for the National Guard families and Department of Paediatric Dentistry at Barts and The London Dental Hospital which is a regional specialist centre for children.

2.1.3 The selection criteria

Patients who had agenesis of at least one permanent tooth, other than third molars, from consanguineous families were recruited. Patients were excluded if they had other syndromic features or systemic diseases, or they were under the age of six since hypodontia diagnosis cannot be confirmed at a young age.

2.1.4 Study population

Sixteen probands were identified comprising 15 families from Saudi and one Pakistani family from London. Together with families, a total of 364 individuals were recruited as described in the following sections.

2.1.4.1 Saudi families

Fifteen children with non-syndromic familial tooth agenesis referred to the paedodontics and orthodontics clinics at the National Guard, King Abdul Aziz Medical City in Riyadh, Saudi

Arabia for treatment were recruited. They and their family members were invited and consented to participate in this study. They were examined and interviewed in the paedodontics and orthodontics clinics. Sixty-four subjects with tooth agenesis, 17 males and 47 females (aged 6-55 years), and their unaffected family members, were included.

2.1.4.2 Pakistani family

Three Pakistani siblings with severe tooth agenesis referred to The Barts and The London Dental Hospital for treatment were recruited. They and their family members were invited and consented to participate in this study. They were examined and interviewed in the paediatric dental clinic.

2.1.5 Data collection

Apart from clinical recordings, retrospective dental information was collected by questionnaire (Appendix A3), and from available dental records, all X- rays performed on the children and their relatives were a routine part of their dental care in Saudi Arabia and United Kingdom and were not merely for the purposes of the research study. Recorded data included the recording of missing teeth, extractions of teeth, ectopic teeth, and tooth agenesis of primary dentition. Congenital absence of tooth was confirmed by clinical and radiograph examinations based on the estimated time of tooth development in the literature (Table 2.1).

Table 2.1: The standard population estimated age for teeth to be visible radiographically and clinically (Nanci and Ten Cate, 2008).

Tooth	Radiographically	Clinically
Central incisors	6 months	6-8 years
Lateral incisors	9-12 months	7-9 years
Lower canines	6 months	9-10 years
Upper canines	6 months	11-12 years
Premolars	2-3 years	10-12 years
First molars	At birth	6 years
Second molars	4 years	11-13 years

2.1.6 Pedigree construction

The pedigrees for the families were constructed based on the clinical examinations and interviews. Pedigree analysis was performed for the families to determine the mode of inheritance of the tooth agenesis phenotype. The individuals with congenitally missing teeth were recorded as affected. The presence of other associated dental anomalies found in the families was recorded separately for affected family members and their first- and second-degree relatives.

2.2 Results

2.2.1 The mode of inheritance in the families

The analysis of the family pedigrees suggested an autosomal recessive mode of inheritance with varying degrees of penetrance in 13 families out of 16 (Figures: 2.1, 2.3, 2.4, 2.5, 2.6, 2.7, 2.8 and 2.10). The other three families segregated the disease in an autosomal dominant (Figure: 2.2) and sporadic mode of inheritance (Figures: 2.9 and 2.11). In this thesis, italic Roman numeral refers to Family, non-italic Roman numeral refers to the generation and the Arabic numeral refers to the member in family.

2.2.2 The phenotype of the dental conditions in the families

In general the features of the phenotype were: one or more premolars missing in four families, one or two upper lateral incisors missing in seven families. The association between premolar and upper lateral tooth agenesis was found in three families. The moderate (three families) and severe (two families) phenotypes in these families involved lower lateral incisors, premolars, second molars and upper and/or lower canines. Interestingly, there was no history of deciduous tooth agenesis among all the families reported in this study (Table 2.2).

Table 2.2: Phenotype summary of the 16 studied families, demonstrate the type of teeth missing and other dental anomalies associated with the tooth agenesis in each family.

Family	No of family members	No of affected members		Type of inheritance	Teeth missing	Impacted canine	Generalised microdontia	Peg shaped lateral	Other dental anomalies
		Female	Male						
I	85	14	8	Autosomal recessive	U lateral & premolars	3	8	3	None
II	9	3	4	Autosomal dominant	Premolars	0	0	0	None
III	8	1	2	Autosomal recessive	U canines, lateral incisors, premolars. L lateral incisors, canines and premolars				Evagination Conical incisors Anterior open bite Incisal notches Taurodont
IV	11	1	2	Autosomal recessive	Premolars	0	0	0	None
V	19	2	0	Autosomal recessive	Premolars	0	0	0	None
VI	19	3	0	Autosomal recessive	Premolars	0	0	2	None
VII	27	5	1	Autosomal recessive	U lateral & premolars	0	5	0	None
VIII	14	2	1	Autosomal recessive	U lateral	0	0	1	Rotated canine
IX	21	2	0	Autosomal recessive	U lateral	0	2	2	None
X	23	4	0	Autosomal recessive	U lateral	0	0	0	Rotated lateral and canine
XI	16	3	1	Autosomal recessive	U lateral	0	3	0	None
XII	20	2	0	Autosomal recessive	U lateral	3	0	0	None
XIII	9	2	0	Autosomal recessive	U lateral	0	2	0	None
XIV	25	1	0	Sporadic	L lateral second premolars U& L molars	1	4	1	Peg shaped laterals
XV	26	2	0	Autosomal recessive	L Lateral	0	0	0	None
XVI	14	1	0	Sporadic	Canine	0	3	1	None
Total	**364**	**48**	**19**						

2.2.2.1 Family I with mild to moderate tooth agenesis of premolars and upper lateral incisor

A large Saudi family with a highly consanguineous pedigree of 85 members included 22 with non-syndromic selective tooth agenesis of the premolars and upper lateral incisors (8 males and 14 females).Six branches in the pedigree were showing first cousin marriage. The disease mutation appeared to be segregating in an autosomal recessive manner, although in some branches a semi-dominant pattern existed with one of the parents and his or her children affected (Figure 2.1). The family showed varying degrees of penetrance in the number and type of teeth affected (Figure 2.1). The phenotypes of the tooth agenesis and other dental anomalies among the affected members are listed in Table 2.3.

Table 2.3: Summary of phenotype of members with tooth agenesis in *Family I.*

Member	Age	Gender	**Phenotype**	
			Type of teeth missing	Other dental anomalies
II:1	55	Male	(2) Upper left lateral incisor & second premolar	None
II:3	48	Male	(1) Lower left .second premolar	None
II:4	31	Female	(2) Lower .second premolars	None
II:5	35	Male	(2) Lower .second premolars	None
II:6	37	Female	(2) Lower .second premolars	None
II:8	27	Male	(2) Lower .second premolars	None
II:14	42	Male	(1) Lower right .second premolar	None
II:16	44	Female	(3) Upper left .second premolar & lower .second premolars	Impacted right canine
II:23	51	Female	(4) .second premolars	Impacted canines
III:19	35	Female	(2) Lower .second premolars	None
III:7	9	Female	(2) Lower .second premolars	None
III:20	33	Female	(2) Lower .second premolars	None
III:22	29	Female	(4) .second premolars	None
III:24	25	Female	(1) Lower left .second premolar	None
III:25	20	Female	(2) Lower .second premolars	None
III:29	41	Female	(2) Lower .second premolars	Impacted left canine & peg shaped lateral incisor
III:30	17	Male	(1) Lower right .second premolar	None
III:40	15	Male	(2) Lower .second premolars	None
III:41	25	Female	(1) Lower right .second premolar	None
III:42	15	Male	(2) Lower .second premolars	None
III:43	21	Female	(2) Lower .second premolars	None
III: 45	19	Female	(2) Lower .second premolars	None

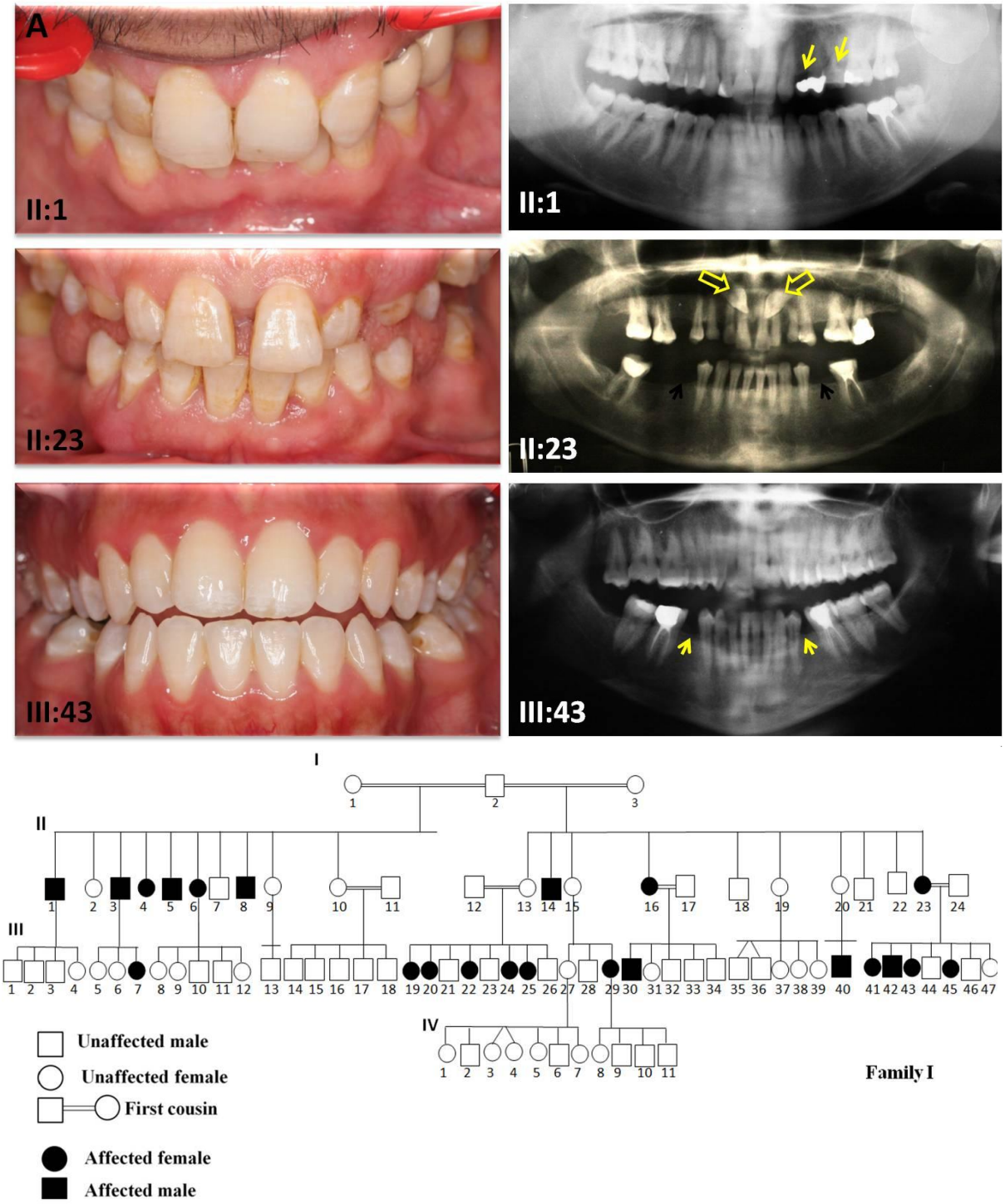


Figure 2.1: *Family I* phenotype.
(A) Clinical photographs (left) and panoramic radiographs (right) of three affected family members (II: 1, II: 23 and III: 43) showing tooth agenesis (solid arrows) and impacted canines (hollow arrows). (B) Pedigree of the family showing the segregation of the condition in the family. The Roman numeral refers to the generation and the Arabic numeral refers to the member in family.

2.2.2.2 Family II with moderate tooth agenesis of premolars and third molars

The parents in the first and second generations of this highly consanguineous Saudi family of non-syndromic premolar and third molar tooth agenesis (Figure 2.2) were first cousins. Out of nine members studied, seven were affected (4 males and 3 females). The tooth agenesis mutation segregated in an autosomal dominant manner (Figure 2.2). Diagnosis of family member I: 2 was based on patient dental records and radiographs (orthopantomogram). The phenotype of the family is illustrated in Table 2.4.

Table 2.4: Summary of phenotype of members with tooth agenesis in *Family II.*

Subject	Age	Gender	Premolars missing	Third molars missing
I-2	72	Female	Upper and lower second premolars	Upper and lower third molars
II-2	48	Male	Upper and lower .second premolars	Lower third molars
III-1	24	Female	Upper and lower .second premolars	Upper and lower third molars
III-2	21	Male	Upper right premolars, upper left .second premolar and lower .second premolars	Lower third molars
III-3	19	Female	Upper and lower .second premolars	Upper and lower third molars
III-4	11	Male	Upper and lower .second premolars	Diagnosis was not confirmed
III-5	9	Male	Upper and lower .second premolars	Diagnosis was not confirmed

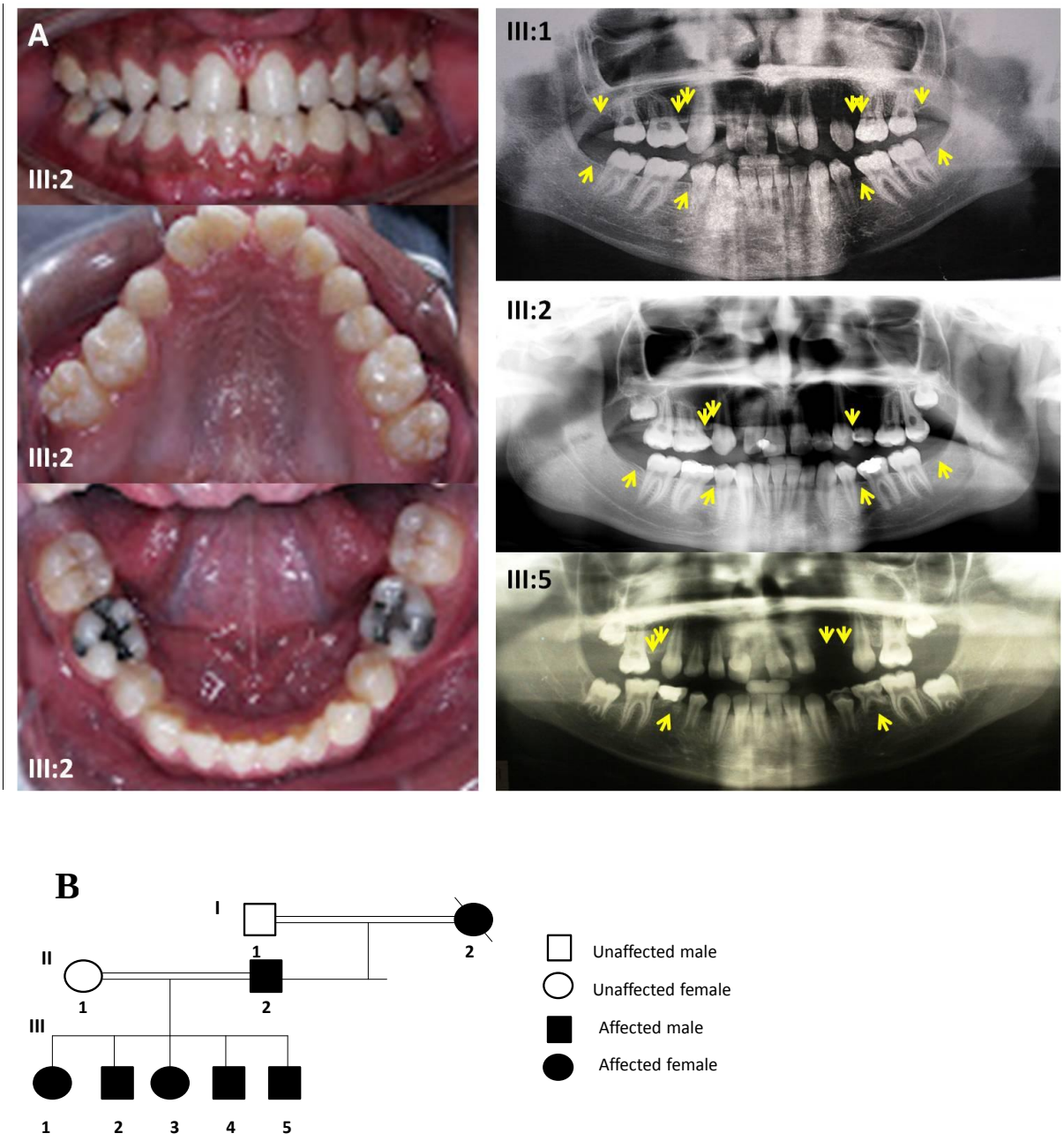


Figure 2.2: *Family II* phenotype.
(A) Clinical photographs of subject III: 2 show the missing teeth and the alignment of the teeth (left) and orthopantomogram radiographs (right) of three affected siblings, III: 1 radiograph shows 10 missing teeth including third molars. III: 2 radiograph shows 7 missing teeth including third molars, and III: 5 radiograph shows 6 missing teeth including third molars. (B) Pedigree of the family showing the segregation of the condition in the family.

2.2.2.3 Family III with severe tooth agenesis of upper and lower premolars, canines and lateral incisors

A Pakistani family presented with severe tooth agenesis that affected three siblings out of six, two males and one female. The severe phenotype in the affected members comprised microdontia and some teeth showing abnormal forms with labial evagination and an incisal notch in the central incisors of varying degree. Also there was evidence of taurodontism in some of their posterior permanent molars. The permanent molars in II: 3 and II: 5 were all single rooted with a thistle shaped pulp (Figure 2.3). Additional clinical features included an anterior open bite due to lack of upper anterior alveolar growth (II: 5 and II: 6) (Table 2.5). All three affected siblings presented with increased amounts of calculus deposits in the lower central incisor region (Figure 2.3).

Table 2.5: Summary of phenotype of members with tooth agenesis in *Family III*

Member	Age	Gender	**Phenotype**	
			Type of permanent teeth missing (excluding 8's)	Other dental anomalies
II:3	**15**	**Male**	Upper canines, second premolars. Lower lateral incisors, canines and premolars	Microdontia Incisal notches Thistle shaped permanent molars
II:5	**10**	**Male**	Upper canines, lateral incisors, left first premolar & second premolars. Lower lateral incisors, canines and premolars	Microdontia Anterior open bite Incisal notches Thistle shaped permanent molars
II:6	**8**	Female	Upper second premolars. Lower lateral incisors, canines and premolars	Microdontia Labial evagination of upper incisor tooth Conical lower incisors Anterior open bite Incisal notches Taurodont and thistle shaped permanent molars

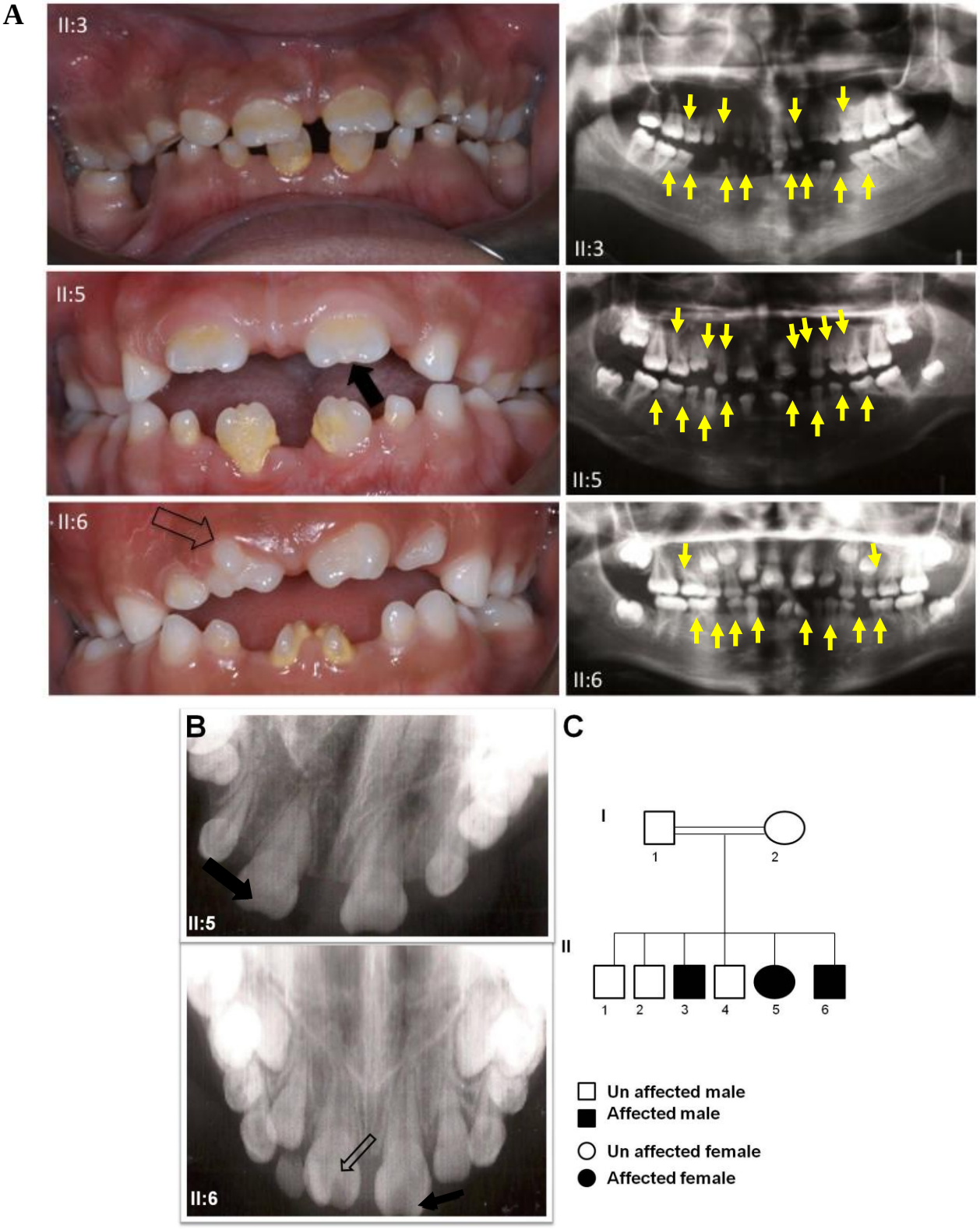


Figure 2.3: *Family III* phenotype
(A) Clinical photographs (left) and panoramic radiographs (right) of the three affected siblings (II: 3, II: 5 and II: 6) showing oligodontia(solid yellow arrow in sibling II: 3, II: 5 and II: 6), severe microdontia. (B) Incisal notches (solid black arrow in sibling II: 5 and II: 6) and evagination of upper incisor tooth (hollow arrow in II: 6). (C) Pedigree of the family showing the segregation of the condition in the family.

2.2.2.4 Families IV, V and VI with mild tooth agenesis of lower second premolars

Families IV, V and *VI* presented with selective tooth agenesis involving only the lower second premolars (Figure 2.4). *Family VI* showed a slightly different phenotype with the association of a peg shaped upper lateral incisor in one family member. All the families segregated the tooth agenesis in an autosomal recessive manner (Figure 2.4). Table 2.6 shows the phenotype of the affected members.

Table 2.6: Summary of phenotype of members with tooth agenesis in *Family (IV, V and VI)*.

Family	Member	Age	Gender	**Phenotype**	
				Type of permanent teeth missing	Other dental anomalies
IV	**II:1**	**43**	**Male**	Lower left second premolar	None
IV	**III:3**	**23**	**Female**	Lower second premolars	None
IV	**III:4**	**20**	**Male**	Lower left second premolar	Rotated lower canines
V	**III:1**	**24**	**Female**	Lower second premolars	Rotated upper right canine
V	**III:7**	**21**	**Female**	Lower second premolars	None
VI	**II:5**	**45**	**Female**	Lower second premolars	Upper pig shaped lateral incisor
VI	**II:6**	**38**	**Female**	Lower second premolars	None
VI	**IV:1**	**10**	**Female**	Lower right second premolar	None

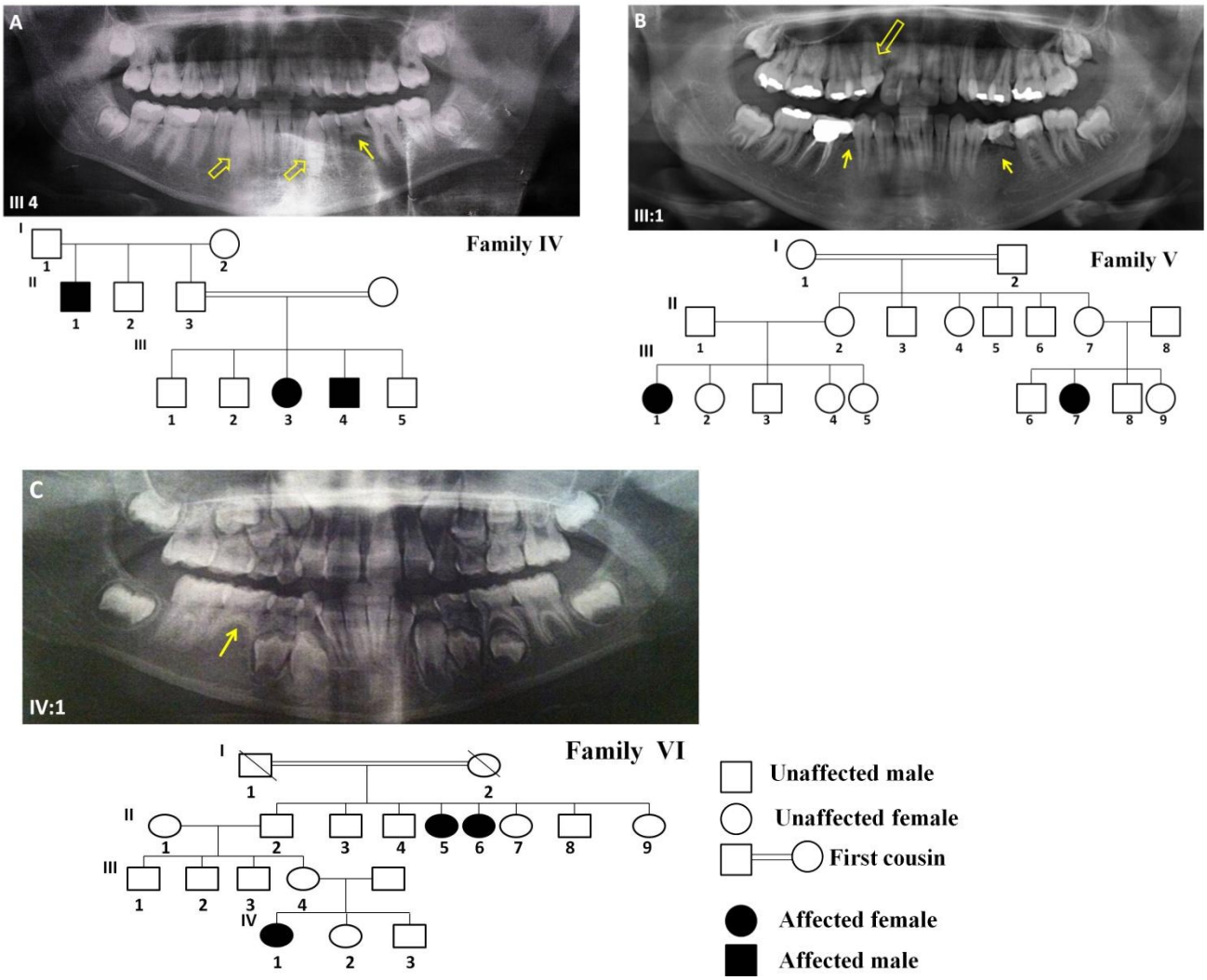


Figure 2.4: *Family (IV, V, and VI)* Lower second premolar tooth agenesis phenotype
(A) ***Family IV***, an orthopantomogram radiograph of one of the affected members (III: 4) showing a missing lower left second premolar (solid arrow) and rotated lower canines (hollow arrows). Notice that all the 3rd molars were present. Pedigree of *Family IV* with 3 affected members segregating the disease in an autosomal recessive manner. (B) ***Family V,*** an orthopantomogram radiograph of one of the affected members (III: 1) showing missing lower second premolars (solid arrows) and rotated upper right canine (hollow arrow). The 3rd molars were present in all affected members. Pedigree of *Family V* with 2 affected members segregating the disease in an autosomal recessive manner. (C) ***Family VI,*** an orthopantomogram radiograph of one of the affected members (IV: 1) showing a missing lower right second premolar (solid arrow). Pedigree of *Family VI* with 3 affected members segregating the disease in an autosomal recessive manner.

2.2.2.5 Family VII with mild to moderate tooth agenesis of upper lateral incisors and premolars

A highly consanguineous Saudi family presented with variable penetrance of the disease mutation and six members affected: II:4, II:7, III:5 and III:7 were missing the upper lateral incisors, III:12 was missing one upper lateral incisor and four second premolars, III:13 was missing one upper second premolar (Table 2.7). The disease mutation segregated in an autosomal recessive pattern. Third molars were present in all the affected members; even III: 12 and III: 13 had all the third molars buds evident in the orthopantomogram radiograph **(**Figure 2.5).

Table 2.7 : Summary of phenotype of members with tooth agenesis in *Family VII*

Member	Age	Gender	**Phenotype**	
			Type of permanent teeth missing	Other dental anomalies
II:4	**53**	**Male**	Upper lateral incisors	None
II:7	**29**	**Female**	Upper lateral incisors	None
III:5	**21**	**Female**	Upper lateral incisors	None
III:7	**24**	**Female**	Upper lateral incisors	None
III:12	**9**	**Female**	Upper left lateral incisor and second premolar & lower second premolars	None
III:13	**8**	**Female**	Upper right second premolar	None

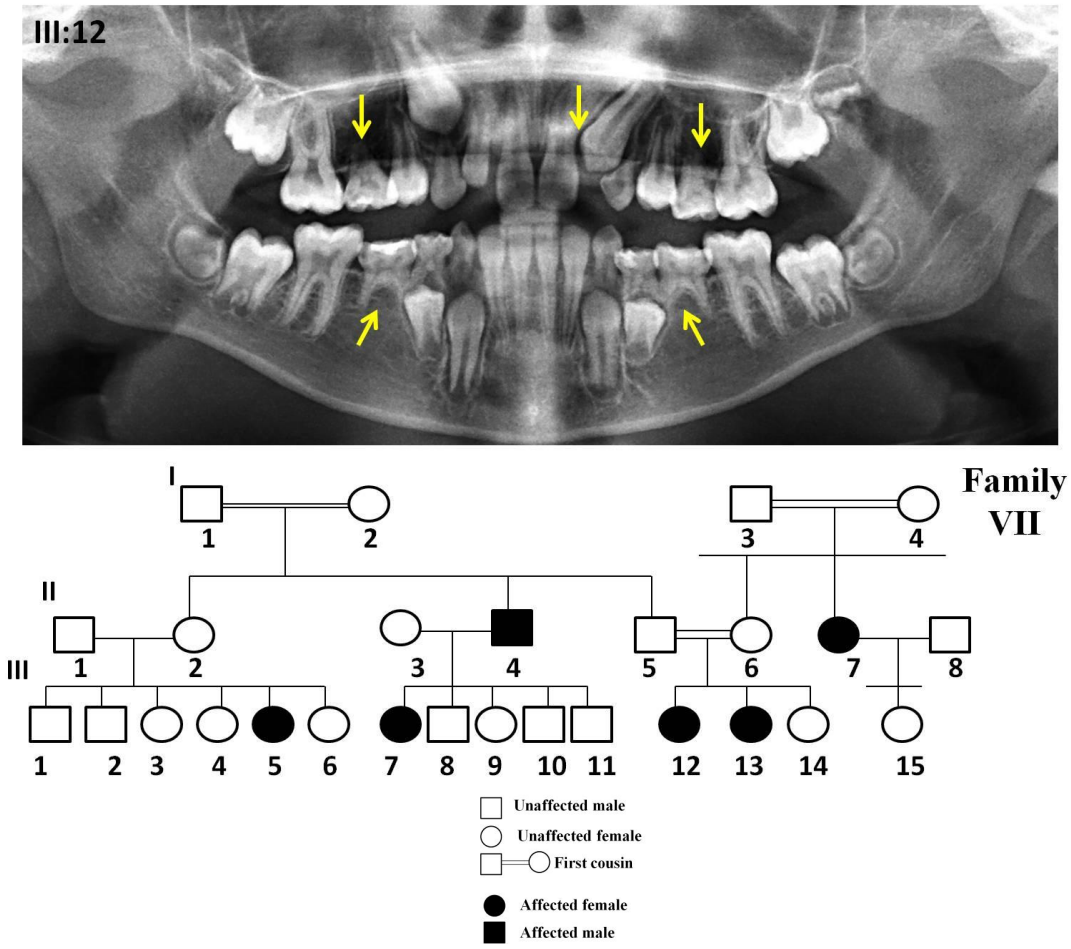


Figure 2.5: *Family VII* phenotype.
(A) An orthopantomogram radiograph of one of the affected members (III: 12) showing a missing upper left lateral incisor and upper and lower second premolars (black arrows). Notice that all the 3rd molar buds were present. (B) Pedigree of the family showing the segregation of the condition in the family, I: 1, I: 2, I: 3 and 1:4 were from the same tribe.

2.2.2.6 Family VIII with mild tooth agenesis of the upper lateral incisors and isolated cleft lip and/or palate in some of those affected

A Saudi family presented with tooth agenesis of the upper right lateral incisor. Three members were affected out of 14. One of the two affected females had a right cleft lip and palate and agenesis of the upper right lateral incisor, the other female was missing the same tooth, the upper right lateral incisor, but with only a right cleft lip. The affected male was missing only the upper right lateral incisor (Table 2.8). The third molar buds were present in all the affected (Figure 2.6).

Table 2.8: Summary of phenotype of members with tooth agenesis in *Family VIII*

Member	Age	Gender	Phenotype	
			Type of permanent teeth missing	Other dental anomalies
III:1	**14**	**Male**	Upper right lateral incisor	Rotated right upper canine, lower canine and lower first premolar
III:3	**8**	**Female**	Upper right lateral incisor	Right cleft lip and palate
III:4	**9**	**Female**	Upper right lateral incisor	Right cleft lip

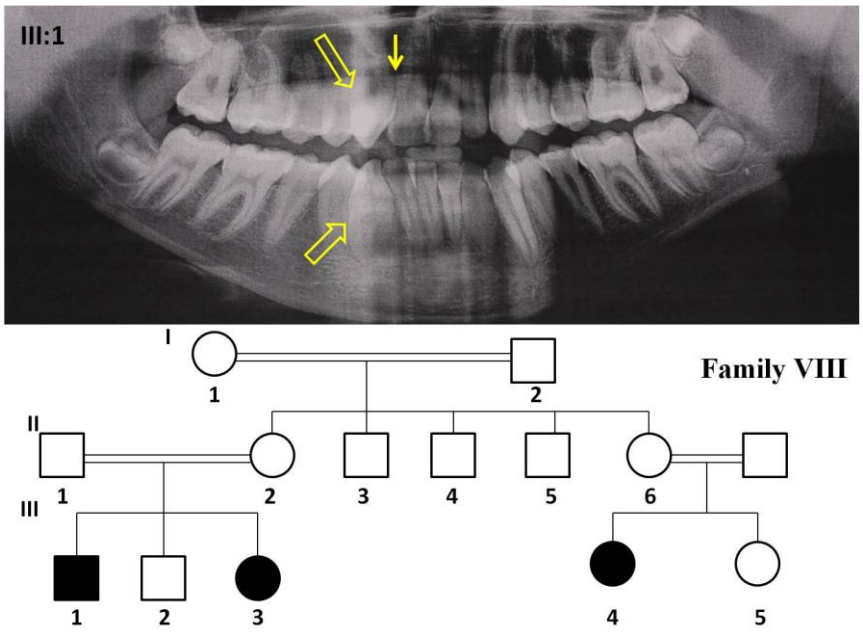


Figure *2.6: Family VIII* phenotype.
(A) An orthopantomogram radiograph of one of the affected members (III:1) showing a missing upper right lateral incisor (solid arrow) and rotated upper right canine, lower canine and lower first premolar (hollow arrows). Notice that all the 3rd molars were present. (B) Pedigree of the family showing segregation of tooth agenesis in the family. III: 3 has right cleft lip and palate. III:4 has right cleft lip.

2.2.2.7 Family IX with moderate tooth agenesis of the upper lateral incisors and third molars

A Saudi family presented with tooth agenesis of the upper lateral incisors inherited in an autosomal recessive manner (Figure. 2.7). Family members were missing two upper lateral incisors (Table 2.9). Third molars were missing in all those affected (Figure 2.7).

Table 2.9: Summary of phenotype of members with tooth agenesis in *Family IX*

Member	Age	Gender	Phenotype	
			Type of permanent teeth missing	Other dental anomalies
III:3	**22**	**Female**	Two upper lateral incisors	Rotated canines
III:5	**19**	**Female**	Two upper lateral incisors	None

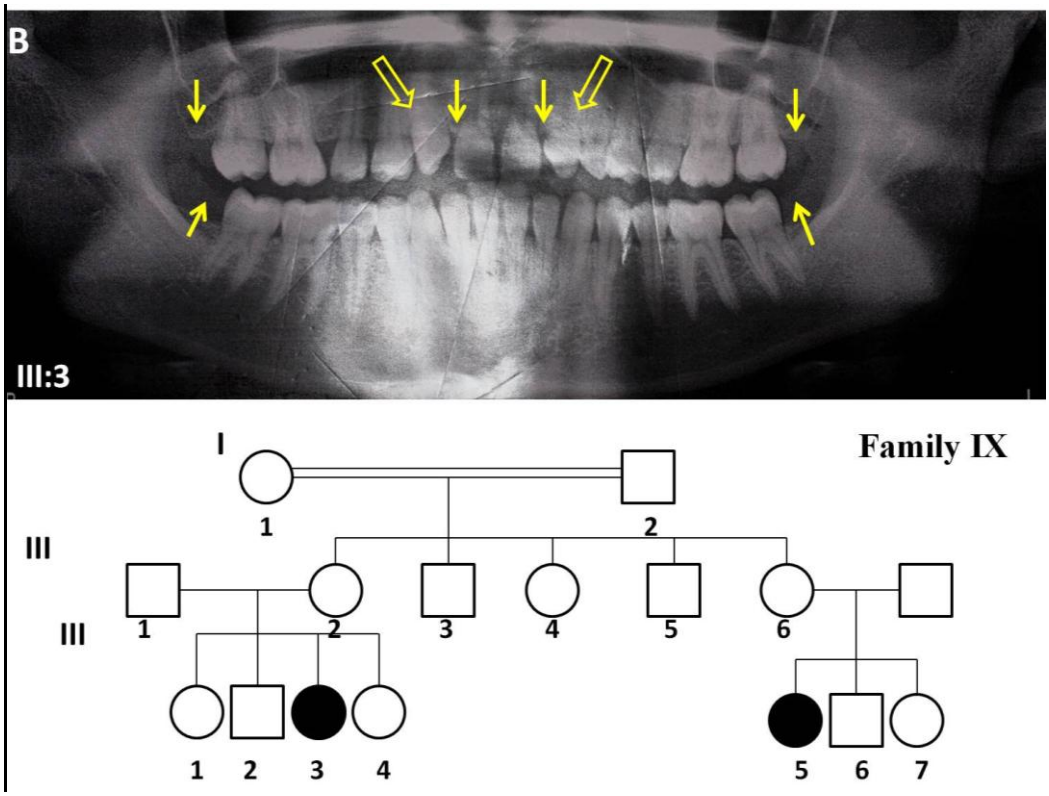


Figure 2.7: *Family IX* phenotype.
(A) An orthopantomogram radiograph of one of the affected members (III: 3) showing a missing upper lateral incisors and rotated upper canines. Notice that all the 3rd molars were missing. (B) Pedigree of the family showing the segregation of the condition in the family.

2.2.2.8 Families X, XI, XII, and XIII with mild tooth agenesis of upper lateral incisors

Families *X*, *XI*, *XII*, and *XIII* presented with agenesis of the upper lateral incisors, inherited in an autosomal recessive pattern in all the families. The phenotype of the members in the families is summarised in Table 2.10 and Figure 2.8. The medical history of *Family XI* showed a high incidence of breast and kidney cancers.

Table 2.10: Summary of phenotypes of *Family X, XI, XII,* and *XIII*

Family	Member	Age	Gender	**Phenotype**	
				Type of permanent teeth missing	Other dental anomalies
X	**II:1**	**46**	**Female**	Two upper lateral incisors	None
X	**III:6**	**40**	**Female**	Two upper lateral incisors	None
X	**III:1**	**17**	**Female**	Upper right lateral incisor	Rotated upper right lateral incisor, lower right canines
X	**III:2**	**15**	**Female**	Two upper lateral incisors	Rotated upper right canine
XI	**II:3**	**51**	**Male**	Upper right lateral incisor	Peg shaped upper lateral incisor
XI	**III:3**	**25**	**Female**	Two upper lateral incisors	None
XI	**III:4**	**24**	**Female**	Two upper lateral incisors	None
XI	**III:8**	**21**	**Female**	Upper right lateral incisor	Peg shaped upper lateral incisor
XII	**IV:1**	**10**	**Female**	Two upper lateral incisors	Impacted canine
XII	**IV:6**	**11**	**Female**	Two upper lateral incisors	Impacted canine
XIII	**II:2**	**39**	**Female**	Two upper lateral incisors	Generalised microdontia
XIII	**III:1**	**14**	**Female**	Two upper lateral incisors	Generalised microdontia

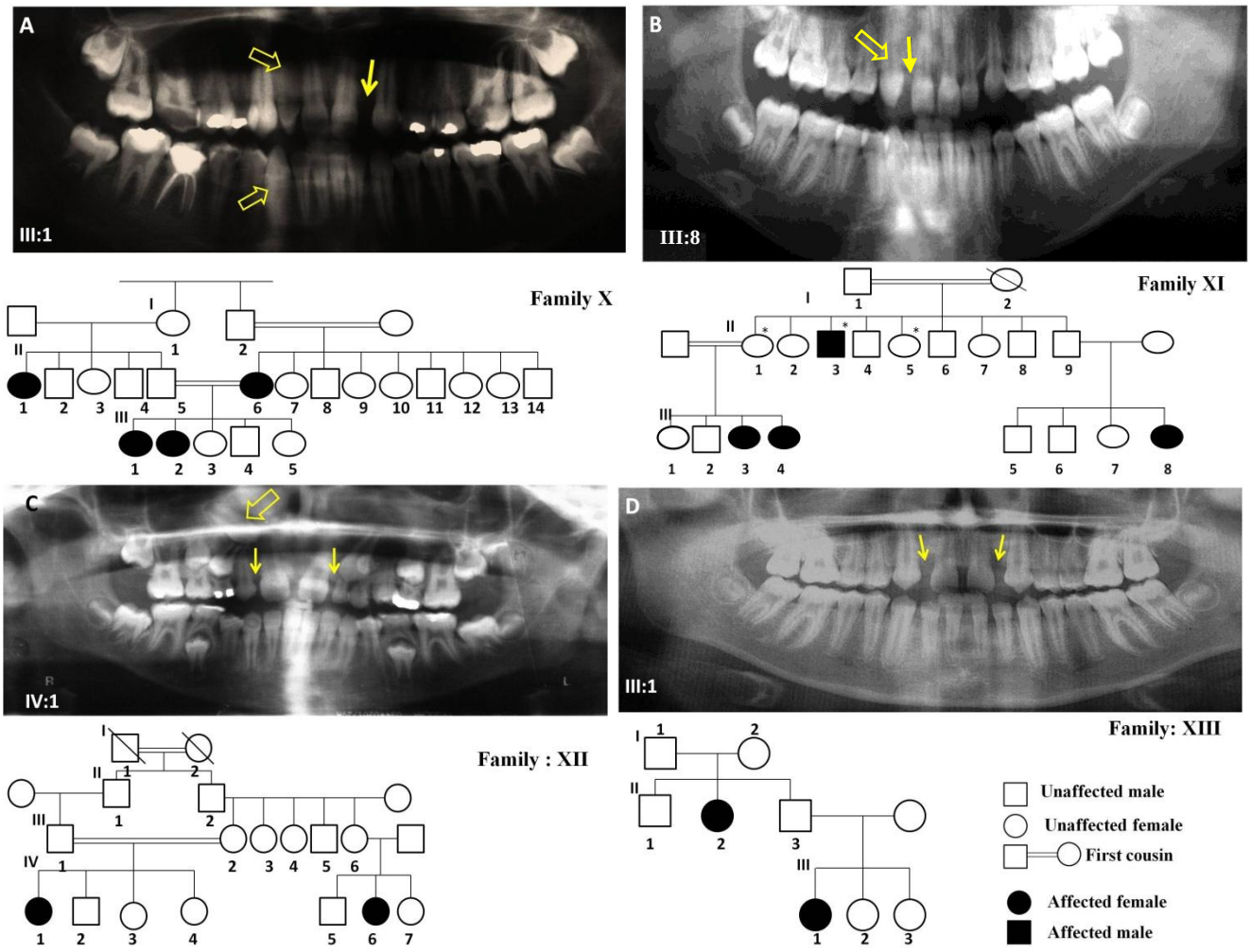


Figure 2.8: *Family X, XI, XII,* and *XIII* upper lateral incisor tooth agenesis phenotype

(A) *Family X*, an orthopantomogram radiograph of one of the affected members (III:1) showing unilateral absence of the upper lateral incisor on the left side (solid arrow) and rotated upper lateral incisor and lower canine on the other side (hollow arrows). Notice that all the 3rd molars were present. Pedigree of *Family X* with 4 affected members segregating the disease in an autosomal recessive manner. (B) *Family XI* an orthopantomogram radiograph of one of the affected members (III: 8) missing the upper right lateral incisor (solid arrow) and showing a rotated upper right canine (hollow arrow). The 3rd molars were present in all the affected members. (C) *Family XII*, an orthopantomogram radiograph of one of the affected members (IV: 1) showing a bilateral absence of the upper lateral incisors (solid arrows) and an impacted upper right canine (hollow arrow). Pedigree of family *XII* with 2 affected members segregating the disease in an autosomal recessive manner. (D) *Family XIII*, an orthopantomogram radiograph of one of the affected members (III: 1) showing a bilateral absence of the upper lateral incisors (solid arrows) and generalised microdontia. Pedigree of *Family XIII* with 2 affected members segregating the disease in an autosomal recessive manner.

2.2.2.9 Family XIV with sporadic severe tooth agenesis of molars premolars and incisors

A Saudi family with non-syndromic tooth agenesis presented with one female member affected with tooth agenesis with no history of any systematic disease or cancer. Two female members had upper peg shaped lateral incisor and impacted canines. The only affected female had a severe type of tooth agenesis and had missing all the second molars, lower second premolars, lower left central incisor and 3 third molars. (Figure 2.9 and Table 2.11).The pedigree of the family shows that tooth agenesis was presented as a sporadic case with no other affected members.

Table 2.11: Summary of phenotype of members with tooth agenesis in *Family XIV*

Member	Age	Gender	**Phenotype**	
			Type of permanent teeth missing	Other dental anomalies
II:2	**47**	**Female**	None	Peg shaped upper lateral incisor
III:7	**12**	**Female**	Second molars, lower second premolars, lower left central incisor, upper and lower third molars *.	Rotated upper left canine
III:8	**21**	**Female**	None	Impacted upper canines

* The diagnosis of the third molars could not be confirmed as the age of the child was only 12 and late formation of the third molars could occur.

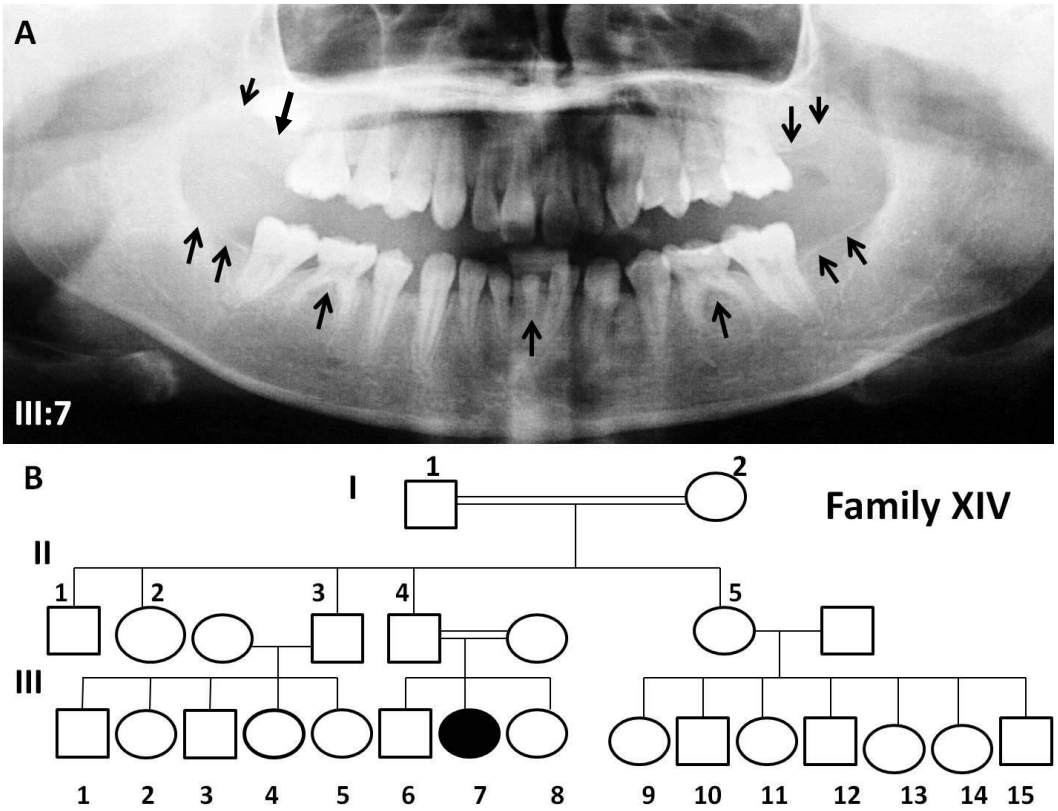


Figure 2.9: ***Family XIV*** phenotype
(A) An orthopantomogram radiograph of one of the affected members (III:7) showing agenesis of the upper second molars, lower second molars and second premolars, lower left central incisor, upper and lower third molars (black arrows). (B) Family's pedigree showing only one affected in the family (sporadic). The Roman numeral refers to the generation and the Arabic numeral refers to the member in the family.

2.2.2.10 Family XV with mild tooth agenesis of lower lateral incisor

A Saudi family had missing lower lateral incisors in an autosomal recessive manner (Figure 2.10). Out of 26 members only, two females (II: 5, IV: 3) were affected, and had the same phenotype, *i.e.* agenesis of two teeth (lower lateral incisors). The presence of the third molars in II: 5 was confirmed. There was no history of any syndrome or, cancer recorded in this family.

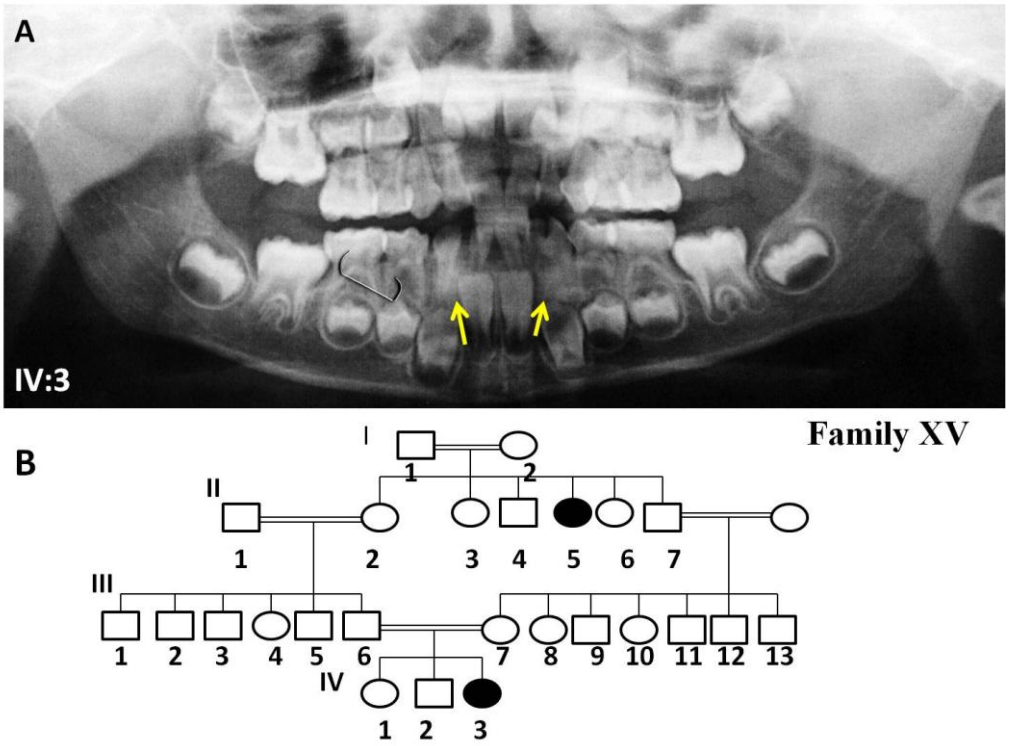


Figure 2.10: *Family XV* phenotype
(A) An orthopantomogram radiograph of one of the affected members (IV: 3) missing lower lateral incisors. (B) Pedigree of the family showing the segregation of the condition in an autosomal recessive in the family. The Roman numeral refers to the generation and the Arabic numeral refers to the member in the family.

2.2.2.11 Family XVI with mild upper canine agenesis

A Saudi family with only one female affected was presented with sporadic tooth agenesis of the upper right canine (other missing teeth were extracted), peg shaped upper laterals and generalised microdontia (Figure 2.11). Fourteen members of the family were checked clinically and radiographically, none were missing any of their teeth, but three males showed generalised microdontia and spaced teeth.

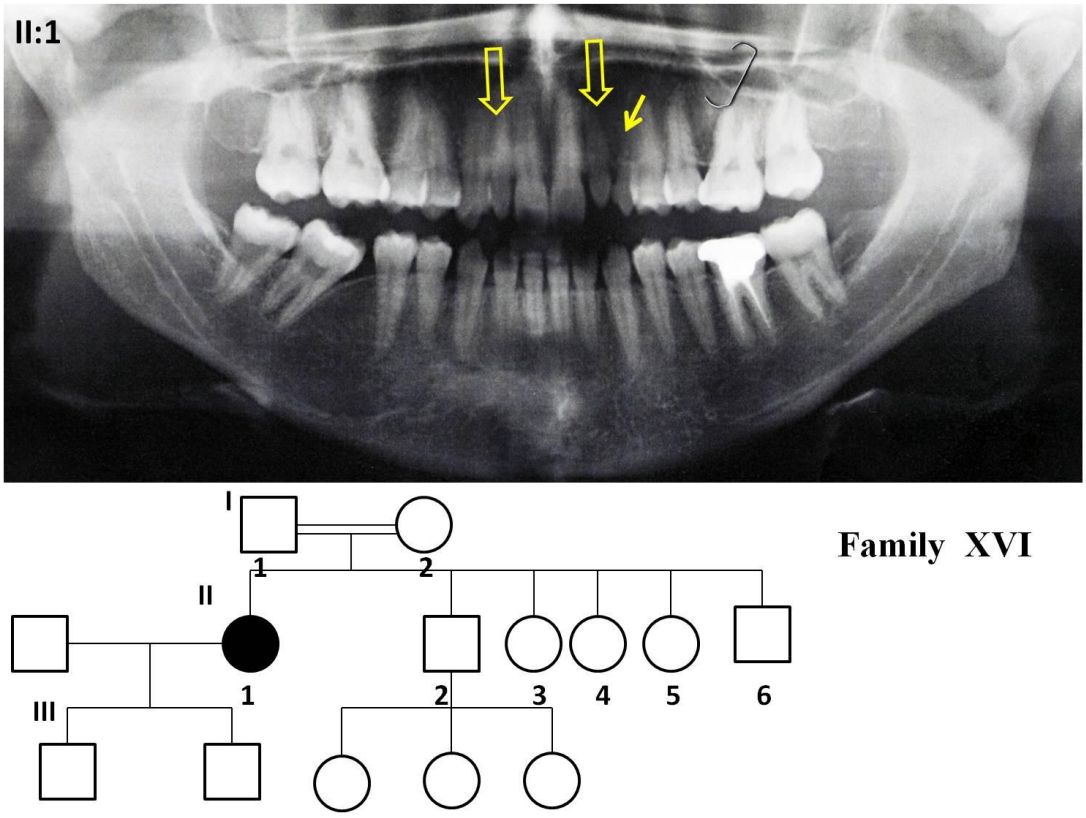


Figure 2.11: *Family XVI* phenotype.
(A) An orthopantomogram radiograph of II: 1 showing a missing upper left canine (solid arrow), upper third molars, retained deciduous upper canines, peg shaped upper lateral incisor (hollow arrows) and generalised microdontia. (B) Pedigree of the family showing the sporadic mode of the condition in the family. I: 1, II: 2 and II: 6 presented with microdontia. The Roman numeral refers to the generation and the Arabic numeral refers to the member in the family.

2.3 Discussion

2.3.1 Classification of tooth agenesis

With the advancement of gene mapping in the last decade, the molecular basis of tooth development and tooth agenesis have been explored. Some genes have been identified that have causative influence to the clinical phenotypes. However, there are phenotypes that cannot be attributed to those discovered genes. Therefore the aim of this study was to characterise the clinical features of each family with tooth agenesis enrolled in this study, classify them on the basis of type of teeth involved, mode of inheritance and degree of severity, in order to link the phenotype of each family to the responsible gene determined by molecular analysis.

Currently the most widely accepted and used definition and classification of tooth agenesis is congenital missing of one or more deciduous or permanent teeth that are resulted from premature cessation of tooth development during its early stages. The term hypodontia is used if only a few teeth are missing (six or less excluding the third molars), and oligodontia if more than six teeth are missing excluding the third molar. However this classical classification was not used in this thesis as these terminologies have been described as limited (Nieminen, 2009) because they do not reflect features of the phenotype that might be significant and specific to a certain gene. For example, the premolar and third molar agenesis phenotype in an autosomal mode of inheritance has been linked to mutations in the MSX1 gene in several case reports (Mostowska *et al.*, 2012). Hence, the type of teeth missing and the mode of inheritance are more likely to be the key, rather than the number of missing teeth, in the classification of tooth agenesis if one wants to find the link to the molecular basis of each condition. The number of teeth missing might reflect the severity of the mutation or the amount of damage in function or changes in protein(s) but is unlikely to lead to the

responsible gene. The agenesis of tooth type and the phenotype was differ depending either on the identity of the mutated gene or the type of mutation, therefore these important pieces of information should be included in the classification of tooth agenesis. At present, this information has not been available for clinicians in the first screening for diagnosis; there is a clear need for recording more clinical information about the type of tooth agenesis to be added to the classification system to aid investigation in molecular level. Clinical classifications need to include the missing tooth type, and the mode of inheritance so that the genetic defect can be identified and patients are more informed.

In this thesis, a novel way of describing non-syndromic tooth agenesis is proposed (Figure 2.12). Firstly, tooth agenesis was divided into three main groups, sporadic, X linked and autosomal (with dominant and recessive subgroups), either in permanent dentition or in both dentitions. Secondly, the severity of tooth agenesis was measured according to the number of tooth types involved rather than the number of teeth. For example, it was proposed that tooth agenesis which involves only premolars should be called mild type, and tooth agenesis which involves only premolars and molars including the third molars could be called moderate type. The third molars have been included in this classification as they may enhance understanding of the impact of the causative mutation, giving important information about the continuity of the effect of the function of the mutated gene in the dentition, and reflecting the severity of the phenotype as well. Tooth agenesis that involves more than two tooth types could be defined as severe type. Thirdly, each subgroups could be subdivided into upper (maxillary), lower (mandibular) or ‘both arches’ (UL) tooth agenesis. Fourthly, the convention should be to state the most affected tooth type first. With this classification, the clinicians will have more information to handle the dental condition and this might help in future investigation of molecular analysis. Using this method of classification, as summarised in Figure 2.12, a

resultant categorisation tree figure was established to enable investigation for the link of the phenotype in these reported cases to their molecular data.

In this study, the 16 families were classified as follows. Two families had sporadic tooth agenesis: one had the mild type of missing upper canine only (*Family XVI*) and another had a severe type of missing molars, lower premolars and incisors (*Family XIV*). There were 14 autosomal families subdivided into one family with dominant moderate tooth agenesis of the upper and lower premolars and third molars (*Family II*), and 13 recessive families. These 13 families were subdivided as follows: one recessive severe tooth agenesis of the upper and lower premolars, canines and lateral incisors (*Family III*); three with recessive mild tooth agenesis of the lower second premolars (families *IV, V* and *VI*); five families with mild tooth agenesis of the upper lateral incisors (families *VIII, X, XI, XII,* and *XIII*; *Family VIII* also had a history of cleft lip and/or palate, and *Family XI* had a history of breast and kidney cancer); one family with mild phenotype of lower lateral incisors (Family XV); three families with recessive mild/moderate tooth agenesis (*Family I* and *Family VII* with agenesis of the second premolars and upper laterals and *Family IX* with agenesis of the upper lateral incisors and third molars) as shown in Figure 2.7. By classifying the families by their mode of inheritance and type of missing teeth, family selection for molecular analysis to explore the wide spectrum of non-syndromic tooth agenesis was facilitated. This enabled a focus on the families with the same mode of inheritance with a similar phenotype, and for exome sequencing.

Most of the families had the most common phenotype of one or two teeth missing, there were some other families showing less common phenotypes which involved the most reported stable tooth type such as canines. The mild phenotype of upper lateral incisor or lower second

premolar tooth agenesis was the most common phenotype among the sixteen families. Six families were showing the lower second premolars phenotype and eight were showing the upper lateral incisor phenotype. This is consistent with what has been found about the Saudi population (Al-Emran, 1990; Salama and Abdel- Megid, 1994), and with what has been reported by Polder *et al.*, (2004) in their meta-analysis study, revealing that in Caucasian populations the most commonly missing teeth were the lower second premolars, followed by the upper lateral incisor and upper second premolar. The second premolars and upper lateral incisors comprised about 85% of all other missing teeth.

2.3.2 Association of dental anomalies and tooth agenesis

Several studies have shown an association of dental anomalies with tooth agenesis. In *Family I,* three affected females with second premolar agenesis out of 12 had palatally impacted canines and one of them also had peg shaped lateral incisors. In addition, *Family XII* had agenesis of the upper lateral incisors and palatally impacted canine. This finding therefore supports the hypothesis that the palatally impacted canines are part of the spectrum of tooth agenesis, mainly related to the mild type of upper lateral incisor or lower second premolar agenesis (Svinhufvud *et al.*, 1988; Bjerklin *et al,.* 1992; Pirinen *et al.*, 1996). It also supports previous findings about the associations between tooth agenesis, upper peg shaped laterals and palatally impacted canines (Bass, 1967). This new finding has very significant clinical implications as canine impaction might complicate dental treatment by causing root resorption of the adjacent teeth. Therefore in the assessment of tooth agenesis, clinicians should be aware of the high risk of the associated canine impaction; thus early screening for impacted canines in patients with familial tooth agenesis and their relatives is recommended, especially when the normal eruption time has been exceeded.

Family III presented with a phenotype of severe tooth agenesis and other dental anomalies expressing the severity of the molecular changes in the dentition related to the mutated gene. Ahmad *et al.,* (1998) reported an association of the severe type of autosomal recessive tooth agenesis with dental anomalies such as microdontia, crown malformation, enamel hypoplasia and hypocalcification (Ahmad *et al.,* 1998). Bloch-Zupan *et al.,* (2011) presented a severe phenotype that was similar to the phenotype reported by Ahmad *et al.* and to the one found in *Family III* described in this thesis (Bloch-Zupan *et al.,* 2011). However there were slight variations between the reported cases. For example, the presence of macrodontia in the family by Bloch-Zupan *et al.,* (2011), and enamel hypoplasia in the family by Ahmad *et al.* (1998), were not found in *Family III* in the present study. *Family III* also presented with taurodontism which is consistent with the family reported by Arte *et al.,* (2001). *Family VIII*, *IX* and *X* presented with agenesis of the upper lateral incisors and rotated canines and/or first premolar; these findings are consistent with Baccetti's observations about the association between agenesis of the upper lateral incisors and rotation of other teeth (Baccetti, 1998a). Many researchers have observed a peg shaped upper lateral incisor in association with tooth agenesis (Alvesalo and Portin, 1969; Al-Emran, 1990; Salama and Abdel-Megid, 1994; Baccetti, 1998b; Arte *et al.,* 2001); this observation is consistent with the phenotype of several families in the present study such as *Family I, IX* and *XIV*. Alvesalo and Portin (1969) proposed that aberrations and agenesis of lateral incisors are different expressions of the same genetic trait as tooth agenesis (Alvesalo and Portin, 1969). This suggestion cannot be verified unless the research into the molecular basis of tooth agenesis succeeds in identifying the modifying genes and discovering their functional roles in affecting the expression of the main gene for each tooth agenesis condition. This is now an achievable goal for future research, especially with advances and developments in high throughput sequence capture methods and next generation sequencing technologies.

An interesting observation was made in *Family VIII* who presented with a history of cleft lip and/or palate. The proband presented with unilateral tooth agenesis of the upper right lateral incisors, which seems to be the minor phenotype of the mutation responsible for the right cleft lip and palate. His sister presented with right cleft lip and palate and agenesis of the upper right lateral incisor; and his cousin had a right cleft lip and right agenesis of the upper right lateral incisor. Van den Booggard *et al.,* (2000) described an association between tooth agenesis and oral facial clefts caused by mutation in the *MSX1* gene (van den Boogaard *et al.,* 2000). However, several other case reports on patients with oral clefts caused by *MSX1* gene mutation did not find any dental anomalies (Jezewski *et al.,* 2003; Suzuki *et al.,* 2004). This finding in *Family VIII* suggests that tooth agenesis might occur as a minor phenotype of the cleft lip and palate trait. This family was found to have right lateral incisors agenesis, contrary to Shapira *et al'*s report that tooth agenesis was found more on the left side, as it was the most affected side in clefting (Shapira *et al.,* 2000b). The distribution of tooth agenesis in the current study was similar between the left and right sides, and upper and lower arches, which supports Polder *et al'*s meta-analysis (Polder *et al.,* 2004). It has been reported that the most stable permanent teeth are the upper central incisors (prevalence of missing = 0.016%), and lower first molars and canines (0.03%) (Nieminen, 2009). This was true of the upper central incisors and first molars but not the canines in the current study, as two families out of the 16 had missing canines. Interestingly, the lower first premolars were present in 15 of the 16 studied families, showing the lower first premolar is a very stable tooth type, as reported being present in most reported cases (Table 2.2).

It is interesting to note that in all sixteen families of this study, some of the dental anomalies that were mentioned in the literature such as infra-occlusion of primary molar and enamel hypoplasia, hypocalsifications were not found. These findings may be explained by the fact that twelve families out of sixteen were showing the mild phenotype.

2.3.3 Mode of Inheritance

There is a wide genetic heterogeneity on tooth agenesis in the present study suggesting that the common types of tooth agenesis may have different modes of inheritance. However, in severe tooth agenesis cases, the autosomal dominant mode of inheritance with reduced penetrance and variable expression is the most common mode of inheritance (Arte *et al*., 2001).

2.4 Summary

The results of the present study support previously reported observations on other dental anomalies associated with tooth agenesis and provide evidence about different phenotypes of tooth agenesis with different patterns of inheritance in consanguineous families. Twelve of the sixteen reported families presented with the mild phenotype of one tooth type agenesis of lower second premolars and/or upper lateral incisors (excluding the third molars), supporting that these are the most common phenotype as reported in the literature.

Thirteen families, segregated by the most common phenotype of tooth agenesis, had an autosomal recessive trait; this may have resulted from the consanguinity of the families, but there is also a high incidence of the mild type of agenesis of the lower second premolars and upper lateral incisors as in Caucasian populations. Although the population used in this study was selected for autosomal recessive inheritance by using consanguineous families, the phenotype, i.e. the pattern of missing teeth, are similar to that found in other populations of non-consanguineous families.

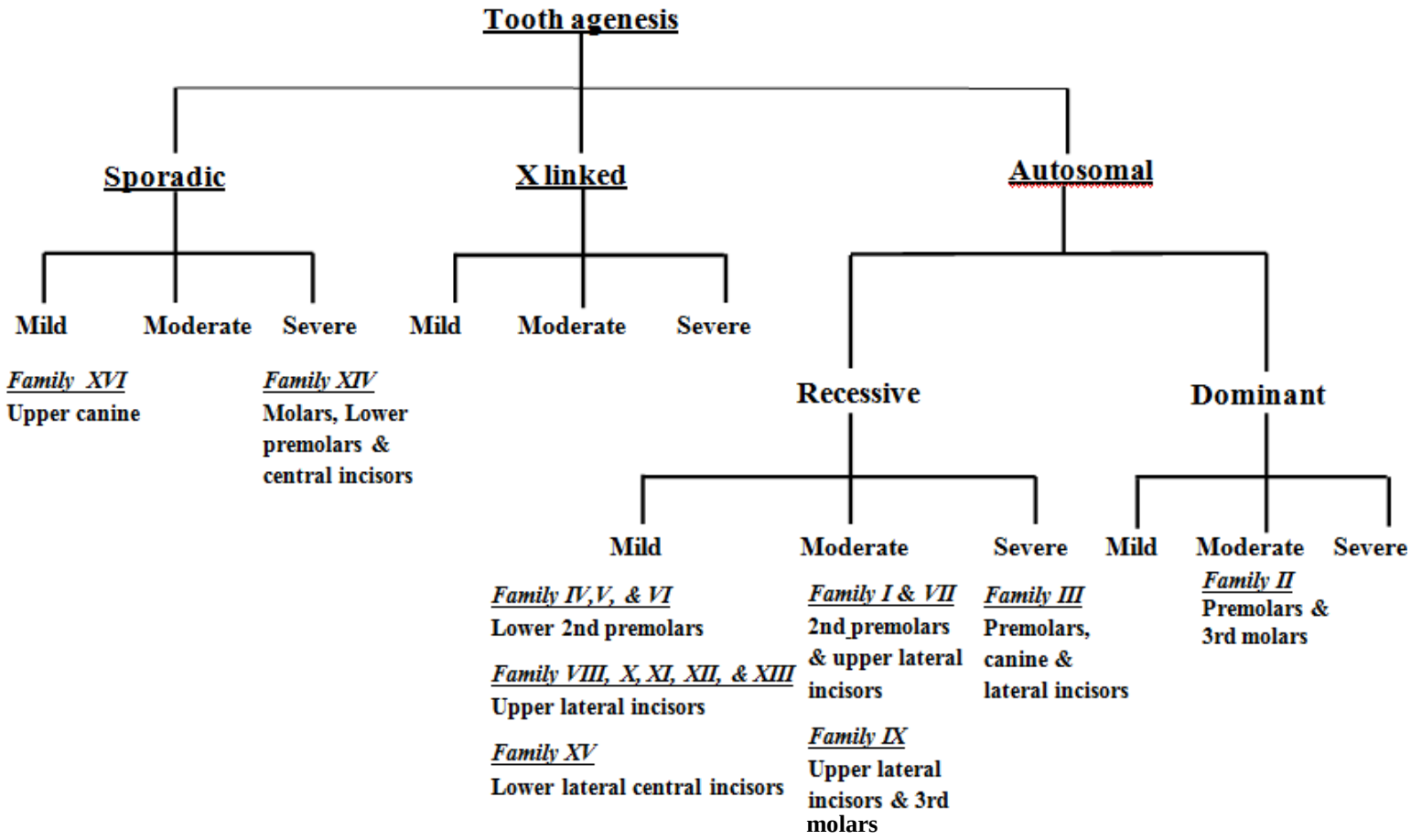


Figure 2.12: Tree figure shows the proposed systematic categorization of the families in this study.

CHAPTER 3: THE GENETIC INVESTIGATIONS TO IDENTIFY THE DISEASED MUTATIONS IN FAMILIES WITH TOOTH AGENESIS

3.1 Background

The identification and characterisation of particular genetic variations that influence tooth agenesis and its different phenotypes is complicated by the influence of other genetic and environmental factors. Thus, many methods and study designed have been used to identify tooth agenesis causal genes and mutations over the last years but few genes have been linked to this dental condition. However the progress towards a full understanding of the genetic basis of tooth agenesis is significantly aided by the revolution in technological development, high- throughput sequence capture methods and next generation sequencing technologies.

Of the 16 families clinically screened and assessed for tooth agenesis, only eight families were identified and selected for the genetic screening. The other eight families, five were excluded because they were different in their phenotypes for example: in *Family VII* the phenotype was different in the two branches of the family among the six affected members, in *Family VIII* the tooth agenesis was associated with cleft lip and palate in the affected females, *Family XV was* missing the lower lateral incisors, *Family XIV* and *Family XVI* were sporadic for which the inheritance pattern could be either autosomal recessive or X linked recessive or even de novo. *Family IX*, *X* and *XII* were excluded because of the lack of unaffected DNA samples. As all the 8 families were from consanguineous populations therefore homozigousity mapping was used initially to locate the regions of the genetic defect underling the tooth agenesis in these families. When this technique failed to identify the causative gene, more sophisticated exome technique was employed. The DNA samples were collected from all the affected individuals (44 individuals). Their parents and other available unaffected siblings (65 individuals, 34 was from *Family I* and 31 from other 7 families, *II, III, IV, V, VI, XI* and *XIII*). Initially families *I* and *II* were selected as a trial, and were screened

for mutations in genes that have been linked before to phenotype similar to their phenotype with regards the tooth type and mode of inheritance (*MSX1 and PAX9* genes). Later *I, IV, V, VI, XI* and *XIII* were selected for homozigosity mapping mainly based on their phenotypes involving only premolar and upper lateral incisor tooth types and on the mode of inheritance which appeared to be autosomal recessive. Due to their consanguinity, large homozygous regions in the families were expected to harbour the disease genes(s). Afterwards whole exome sequencing was performed in three families *Family I, II* and *III.*

The strategy used in this study was aimed to create a systematic approach or criteria in investigating the molecular basis of tooth agenesis by being more specific with regard the tooth agenesis phenotype of each family and the other clinical features that are available for the researcher.

3.2 Materials and methods

3.2.1 Sampling and DNA extraction

The DNA was extracted, depending on the cooperation of the family members, from lymphocytes in patient blood samples only from the affected members and the parents in *Family III* using the QIAmp DNA blood mini/midi kit (QIAGEN, UK) following the manufacturers' instructions. Briefly, buffers in the kit allow samples to be lysed and DNA, which adheres to the column, to be purified (see Appendix B1). DNA from saliva was collected and extracted from the un affected members in *Family III* and from the affected and unaffected members in other Saudi families because of the shipping regulations and difficulties, using Oragene DNA kit (OG-500) (DNA Genotek, Ontario Canada) (Figure 3.1). Briefly, saliva in the collection tube mixes with a preserving fluid and DNA is released. Ethanol precipitation is used to obtain the purified DNA (see Appendix B2).

Figure 3.1: The QIAmp DNA blood mini/midi kit (QIAGEN, UK) and Oragen DNA kit (OG-500) DNA Genotek, Ontario Canada) kits, used for sample collection.

3.2.2 DNA quantification

DNA samples were quantified using a Nanodrop ND-1000 spectrophotometer (Thermo Fisher Scientific, Massachusetts, USA) (Figure 3.2). 1.5µl of sample was loaded on to the Nanodrop pedestal and the optical density (OD) measured. Absorbance at 260 nm was used to calculate the sample concentration. Absorbance at both 260 and 280 nm was used to assess purity. A ratio of 1.8-2.0 was accepted as pure.

Figure 3.2: The Nanodrop ND-1000 spectrophotometer (Thermo Fisher Scientific, Massachusetts, USA) connected to the computer. The picture shows the software used to measure the amount of the DNA.

3.2.3 Primer Design

Primers were designed using the Prime3 website (version 0.4.0) (Rozen &Skaletsky, 2000), accessible at www.frodo.wi.mit.edu/primer3/input.htm, and microsatellite marker primers were designed using the University of California Santa Cruz UCSC Genome Browser website available at www.genome.ucsc.edu/ and the Marshfield Clinic Research Foundation (MCRF) website available at www.marshfieldclinic.org/research/pages/index.aspx. Primers were obtained from sigma-Aldrich, Dorset, UK. Primers used for amplification of genomic DNA are described in Appendix C.

3.2.4 DNA amplification by polymerase chain reaction (PCR)

A standard polymerase chain reaction (PCR) was performed by PCR machine in a 25µl volume compressing: 2.5 µl 10x NH4 buffer, 0.75 µl 50 mM Mg Cl_2, 0.5 µl dNTP mix (200 nM each dNTP), 0.25 µl BIOTAQ (all BIOLINE, London, UK), 0.5 µl 50 µM of each primer and dH_2O with 25 ng template DNA. The PCR enhancing agent betaine and dimethyl sulfoxide (DMSO) 5% or 10% final volume were used to enhance the PCR products if required. Reactions were carried out using a DNA engine tetrad PCR system. The standard amplification cycle is demonstrated in Table 3.1. PCR products were loaded on 1-3% agarose gels depending on the PCR product size.

Table 3.1: PCR amplification cycles

Step	Temperature	Time	Number of cycles
Initialisation	94 ^{o} C	5 min	1
Denaturation	94 ^{o} C	30 s	35
Annealing	* ^{o} C	30 s	
Extension	72 ^{o} C	30 s	
Final extension	72 ^{o} C	10 min	1

*dependent on the primer pair.

3.2.5 Agarose gel electrophoresis

Agarose gel electrophoresis was used to estimate the size of DNA fragments. Agarose (Sigma-Aldrich, Dorset, UK) was dissolved in 1x TBE (Appendix D) in a microwave to give 1-3 gels, cooled and GelRed™ (Biotium, San Francisco, USA) added to the final concentration of 10 µg/ml. Gels were poured into a casting gel tray containing combs, allowed to set and samples were loaded in wells with DNA loading dye, alongside a 1kb plus ladder as required (Invitrogen, California, USA) to allow size comparison. Gels were run in 1xTBE at 100 V. Bands were visualised on a UV transilluminator and photographed (Figure 3.3).

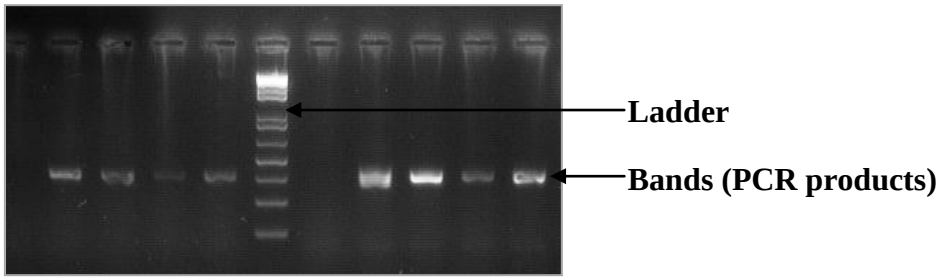


Figure 3.3: A photograph of different PCR product bands with the ladder

3.2.6 PCR purification

PCR products were cleaned prior to sequencing using ExoSap-IT reagent (United States Biochemical, Ohio, USA) following the manufacturer's instructions. This reagent contains the hydrolytic enzymes shrimp alkaline phosphatase and exonuclease I, which degrades surplus dNTPs and primers in the PCR reactions.

3.2.7 Sequencing

Direct sequencing was performed by the cycle sequencing chain termination method. Only one of the PCR primers per reaction is used to copy one strand. The incorporation of fluorescently–labelled dideoxynucleotides into the extending chain of nucleotides causes chain termination and the fluorescent signal is detected and converted into a sequence. Samples were sequenced using the BigDye Terminator version 3.1 Cycle Sequencing Kit (Applied Biosystems, Foster City, CA). 0.5-4 µl of PCR product was added to a reaction containing 1µl primer (10µM) and, 3µl Better Buffer (Microzone, Ottawa, Canada) made up to 11.5 µl with dH_2O. The amplification cycle used is in Table 3.2.

Table 3.2: Sequencing amplification cycles

Step	Temperature	Time	Number of cycles
Denaturation	96 ^oC	1 min	1
	96 ^oC	10 s	25
Annealing	58 ^oC	5 s	
Extension	60 ^oC	4 min	

3.2.8 Post-sequencing clean-up

Sequencing reactions were transferred to plates and precipitated with 2 µl 125mM EDTA and 30 µl ice-cold 100% ethanol, left on ice for 10 min and centrifuged at 4400xg for 20 min. Plates were inverted and centrifuged at 135xg for 10 s to remove the supernatant. The pellet was washed with 135 µl 70% ethanol and centrifuged for 5 min and the supernatant removed as before. Samples were dried in a heating block at 95^oC for 10 s before being loaded on the sequencer. Samples were re-suspended in 10 µl HiDi^{TM} formamide, transferred to an ABI

plate (both Applied Biosystems, Foster City, CA) and denaturised at 95^{o}C for 3 min. Samples were viewed and analysed using Chromas software (Technelysium Pty Ltd).

3.2.9 Sequence analysis

Wild-type reference sequences were obtained from unaffected individuals (controls). Sequences were assembled and compared with wild-type sequences using MultiAlin (Corpet, 1988). The nomenclature for nucleotide positions is indicated by the position from the start of the coding ATG where A=+1 (Den Dunnen & Antonarakis, 2001). Database SNP (dbSNP) (National Centre for Biotechnology Information, Maryland, USA) and the 1000 Genomes Browser (www.browser.1000geneomes.org/index.html) were used to determine if sequence variants discovered were novel or reported SNPs.

Protein sequences between species were aligned and compared using ClustalW (www.ebi.ac.uk/tools/clustalw2/index.html) (Larkin *et al.*, 2007). In this study, sequence variants in dbSNP/1000 Genome Browser are referred to as SNPs and the rs numbers given. Sequence variants which have been reported in control/unaffected individuals, but are not found in dbSNP or the 1000 Genomes Browser, will be referred to as polymorphism.

3.2.10 Illumina DNA Analysis Bead Chips, Infinium® HD assay and homozygosity mapping

Illumina's Bead Array™ technology is based on a random assembly of arrays of 3-micron silica beads in micro wells on either fibre optic bundles or planar silica slides with a uniform spacing of ~5.7 microns. Pre-synthesised oligonucleotides of at least 70 nucleotides in length are used that contain a 25-nucleotide address and a 50-nucleotide gene specific sequence. Each oligonucleotide is fixed to a particular batch of beads which carry more than 100,000 identical oligonucleotides. Beads with different oligonucleotides are collected and then spread across prefabricated microarray microwells designed to accept the bead size. Then the

25-nucleotide address is decoded, letting each bead be identified. Bead Array technology is utilised in Illumina's iScan System for a wide range of DNA and RNA analysis purposes (Strachan and Read 2010) the laboratory steps are illustrated in Figure 3.4.

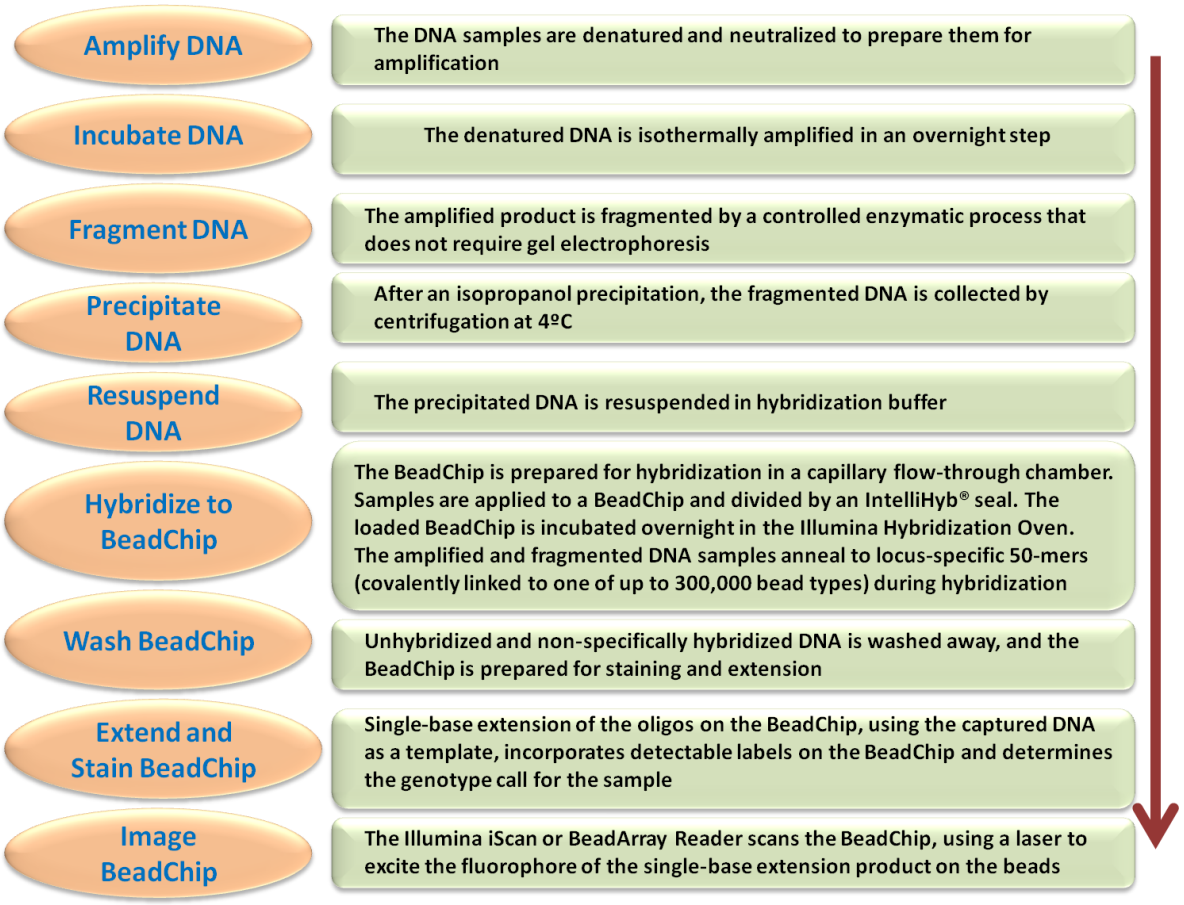


Figure 3.4: Workflow of the laboratory steps of the Illumina Infinium HD BEAD CHIPS as described by the manufacturers.

The DNA of 16 samples of affected members from 6 different families in this study were genotyped using *Illumina Infinium*® high-density arrays single nucleotide polymorphism (300,000 SNPs) by a technician at the Genome Centre, Charterhouse Square, London.

The accuracy of genotypes from the current generation of microarrays is very high and expected to be up to 99.5% (Ritchie *et al.* 2011). The array data were genotyped with genotyping data analysis *Illumina* GenomeStudio™ Genotyping Module v.1 Software.

The genotyping data for all SNPs within a chromosomal region were visualised by the B allele frequency plot and theta as shown in Figure 3.5.

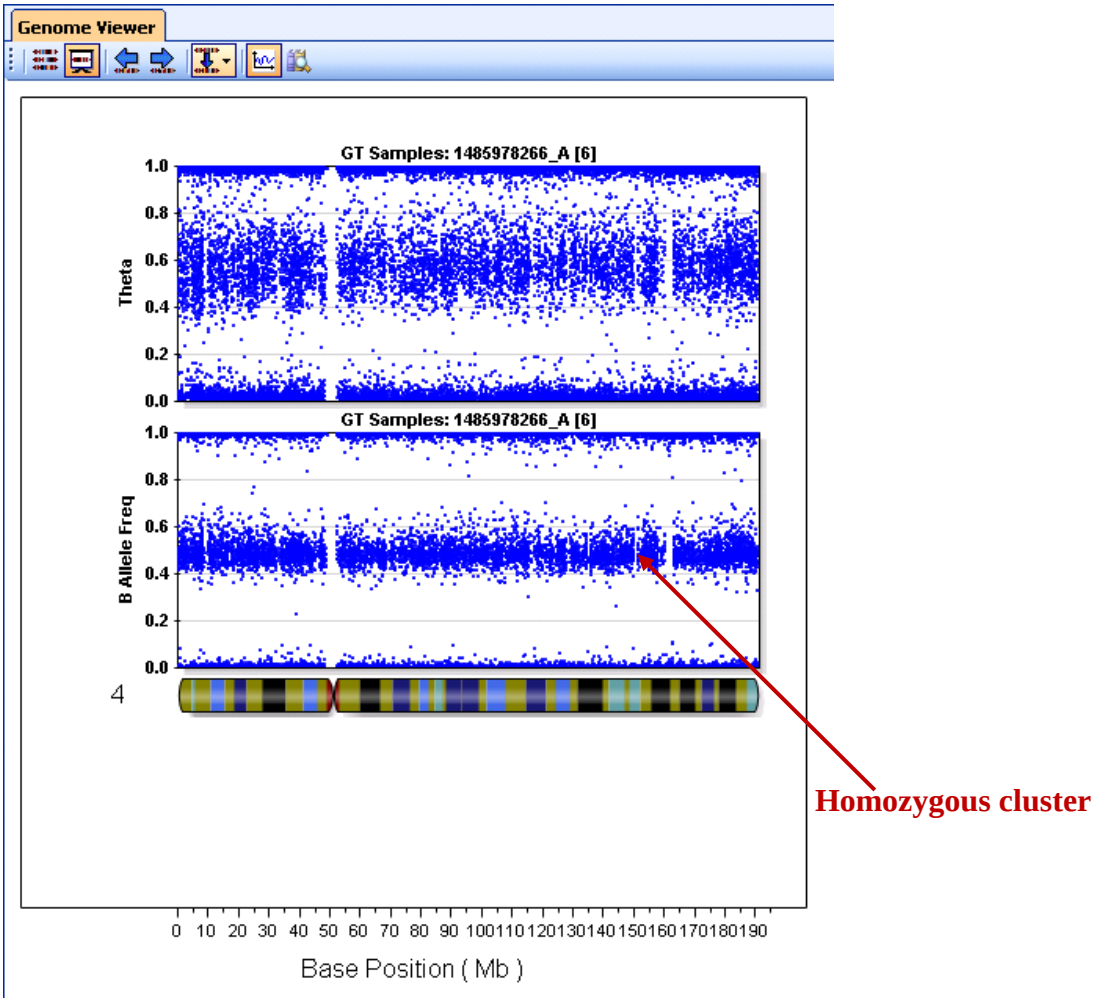


Figure 3.5: Theta values and B Allele frequency for the same sample in chromosome 4.

3.2.11 ABI 3730xl Capilliary Sequencer

Microsatellite marker screens were conducted on affected and unaffected family members carefully selected from the family, using ABI 3730xl Capilliary Sequencer, by a technician at the Genome Centre.

3.2.12 Whole exome sequencing, Illumina HiSeq

Whole exome sequencing was performed on 13 DNA samples of affected members in three families, using a SureSelect Human all Exon 50 Mb kit (Agilent Technologies, Inc., Santa Clara, CA) and sequenced on an Illumina HiSeq. Sequencing reads were aligned to the hg19 build of the human reference genome using the software novoalign (www.novocraft.com) (Figure 3.6). SNP and indel calling were performed using samtools version 0.18 and were annotated using the software ANNOVAR (Wang *et al.*, 2010) by Dr V. Plagnol, Genetics Institute, University College London, UCL. Candidate variants were filtered on the basis of function (as predicted by ANNOVAR), and the 1000 Genomes (www.1000genomes.org) and NHLBI exome sequencing project (http://evs.gs.washington.edu/EVS/) frequencies.

Exome sequencing usually reveals about 2500 variants, after filtering against variants known to exist in the general population. The rare/novel variants that were found to be shared between the affected members in each family were investigated and stratified by their functional class such as frame shift and stop codons, and by existing biological or functional information such as the role of *MSX1*, *PAX9* and *SMOC2* (Figure 3.7).

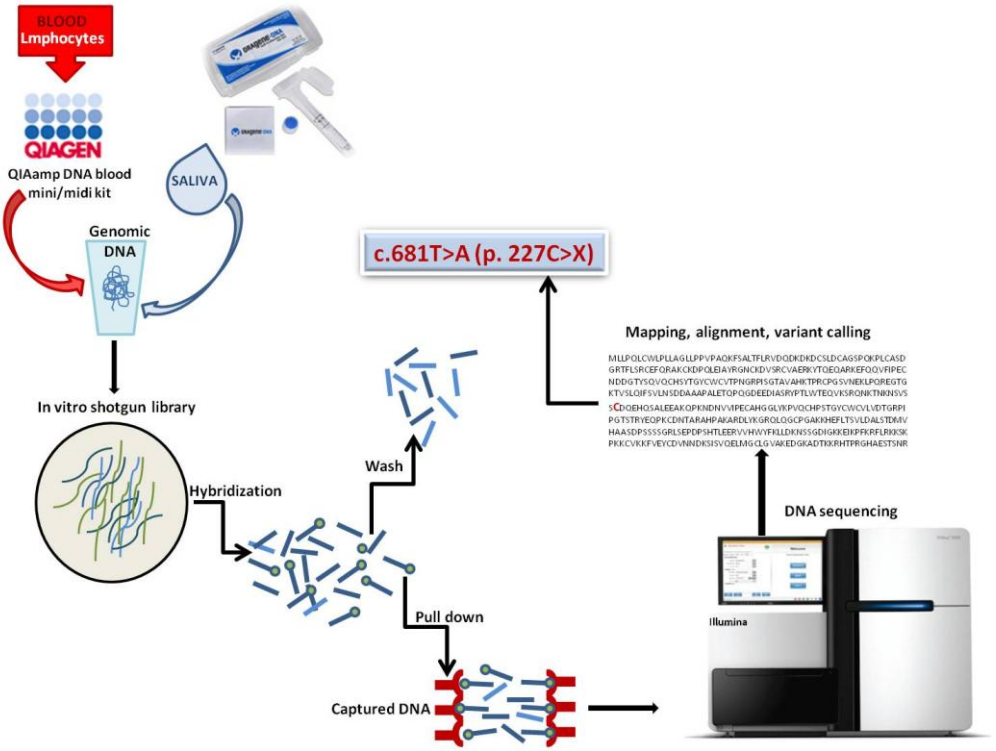


Figure 3.6: Schematic presentation of the whole exome sequencing steps and the Illumina HiSeq

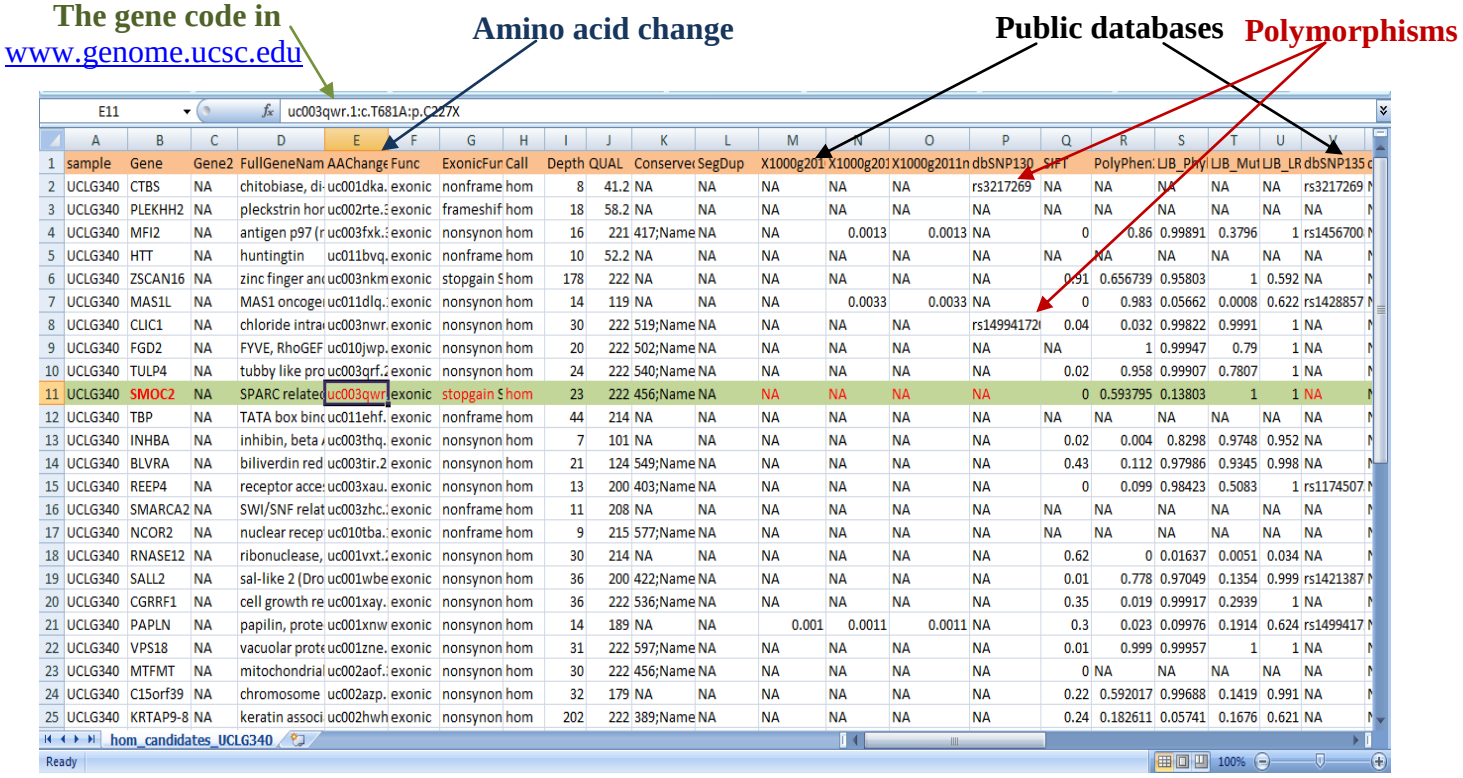


Figure 3.7: An excel sheet shows the exome data presentation after annotation.
The novelty of the variants was assessed by filtering the variants against the polymorphisms available in public data base.

3.3 Linkage analysis and homozygosity mapping

3.3.1 Results

3.3.1.1 MSX1 and PAX 9 genes sequencing

Mutation analysis in *Family I* and *Family II* of MSX1 and PAX9 genes were carried out by direct sequencing of MSX1 two exones and PAX9 5 exons of two affected and one un-affected members in each family; the primers used to amplify the genes are listed in Appendix C1. The direct sequencing of *Family I* and *II* revealed that both families were negative to MSX1 and PAX9 disease mutations.

3.3.1.2 Homozygosity mapping in Family I

Four affected members (II: 23, III: 19, III: 22 and III: 45) were selected from *Family 1* of which two were sisters, one was their aunt and the other was the aunt's daughter. The selection included member of close and maximum distant relation within the family. The four members were genotyped using high density single nucleotide polymorphism (300,000 SNPs) Infinium arrays. MSX1, PAX9, AIXN2 and Wnt10A regions were checked and none of them was homozygous in these samples. Several homozygous regions were found shared among the four samples; these regions are listed in Table 3.3. The largest region was in chromosome 22, it was about 2359234bp, the second largest was in chromosome 19, which was about 1702817bp. This region was rich with zinc finger genes such as ZNF 99, 98, 492, 729, 681, 726 and 730.

Table 3.3: Shared homozygous regions in *Family 1*.

Chromosome	Start rs ID bp	Position	End rs ID bp	Position	SNP no.	Total bp	No. of genes in region
22	rs5752924	28200951	rs9621346	30560185	299	2359234	28
20	rs486344	887161	rs4813941	1076013	42	188852	2
19	rs11669012	22467023	rs3844117	24169840	187	1702817	22
18	rs4890423	38900174	rs627238	39187966	29	287792	1
18	rs9965350	49593834	rs1145303	49910037	46	316203	1
17	rs12603064	26363265	rs2525566	26698980	31	335715	11
16	rs2294618	1754392	rs9928312	1927059	18	172667	11
15	rs1807769	26884934	rs11070422	27148418	22	263484	2
14	rs7144686	102352772	rs4906272	102447074	14	94302	3
12	rs216867	5961261	rs3181301	6216045	49	254784	2
11	rs963090	83356568	rs11233953	83607859	32	272204	1
11	rs938727	84603291	rs6592253	84979990	46	376699	1
9	rs13290574	459626	rs7034698	510886	60	51260	2

3.3.1.3 Homozygosity mapping analysis in families I, IV, V, VI, XI and XIII

Due to the large number of homozygous region shared in the typed individuals from *Family I*, additional family members, and affected individuals from the other families were genotyped A total of 12 affected members were genotyped, four from *Family I* (II:1, II:3, II:14 and III:20), one from *Family IV* (III:4), two from *Family V* (III:1 and III:7), one from *Family VI* (IV:1), two from *Family XI* (III:3 and III:4) and two from *Family XIII* (II:2 and III:1). MSX1, PAX9, AIXN2 and Wnt10A regions were checked but none of them was found to be homozygous in those families. A homozygous region was found shared between three families (*Family I*, *V* and *VI*) in chromosome 22 (29,305,513-29,635,842).

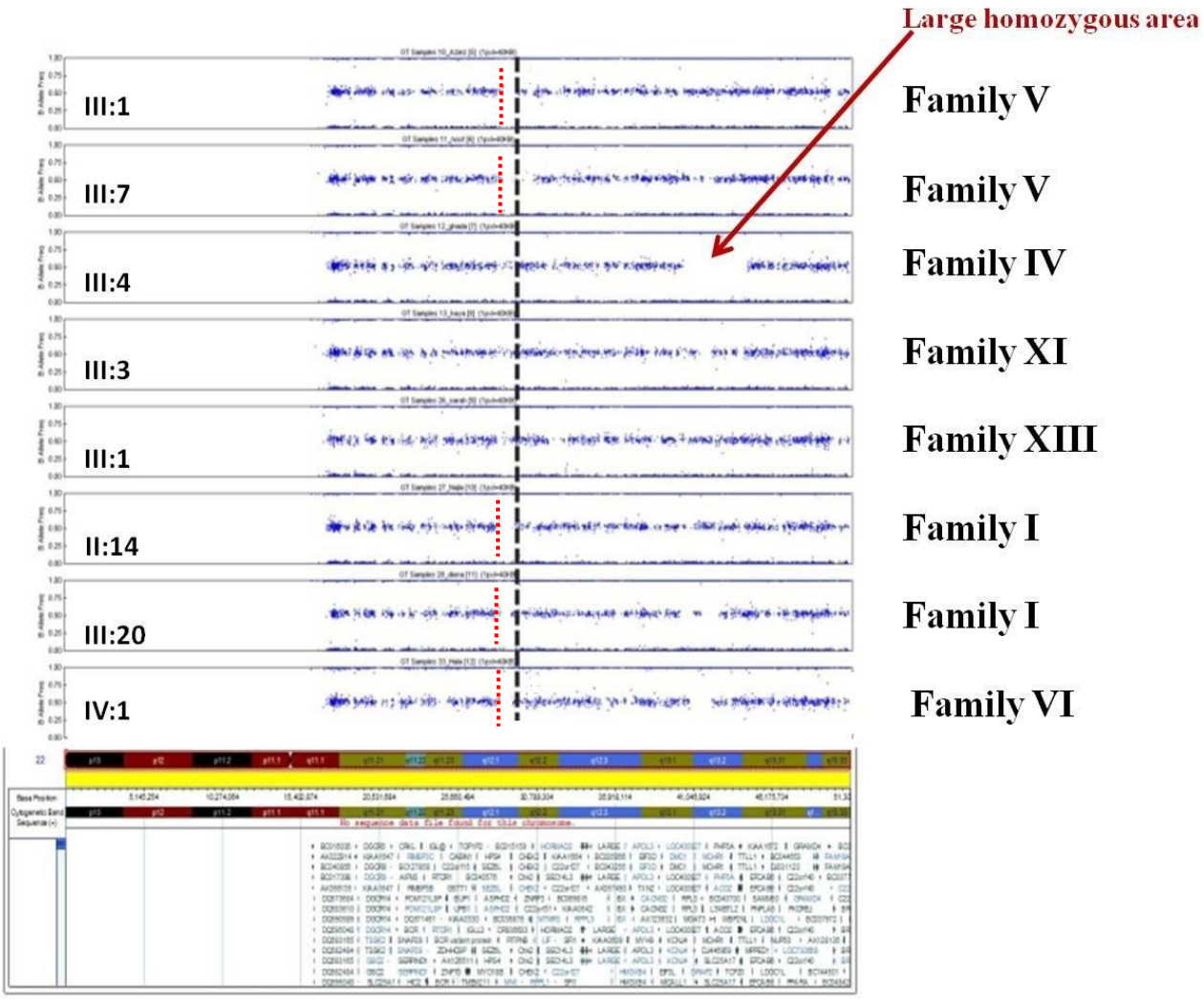


Figure 3.8: Allele frequency view of 8 members of *Family IV, V, VI, IX* and *XIII* showing the *KREMEN1* gene position in the area between the dotted lines, and 3 members in *Family V* and *VI* were sharing this region with Family *I.*

3.3.1.4 KREMEN1 gene Sanger Sequencing

Kringle domain-containing transmembrane protein 1 (KREMEN1) OMIM 609898 was found to be the most likely candidate gene in the homozygous region in chromosome 22 (Figure 3.8).

Mutation analysis of *KREMEN1* gene was carried out by direct sequencing of the 10 exons. Primers used to amplify the *KREMEN1* gene are listed in Appendix C1. The analysis of the 10 exon sequencings did not reveal any mutation in the *KREMEN1* gene that might segregate

in the affected members in the three families. An affected member III: 20 in *Family I* was homozygous for a SNP in *KREMEN1*. A synonymous SNP, c558C>T p.T186T (TMP_ESP_22_29521331) (Figure 3.9).

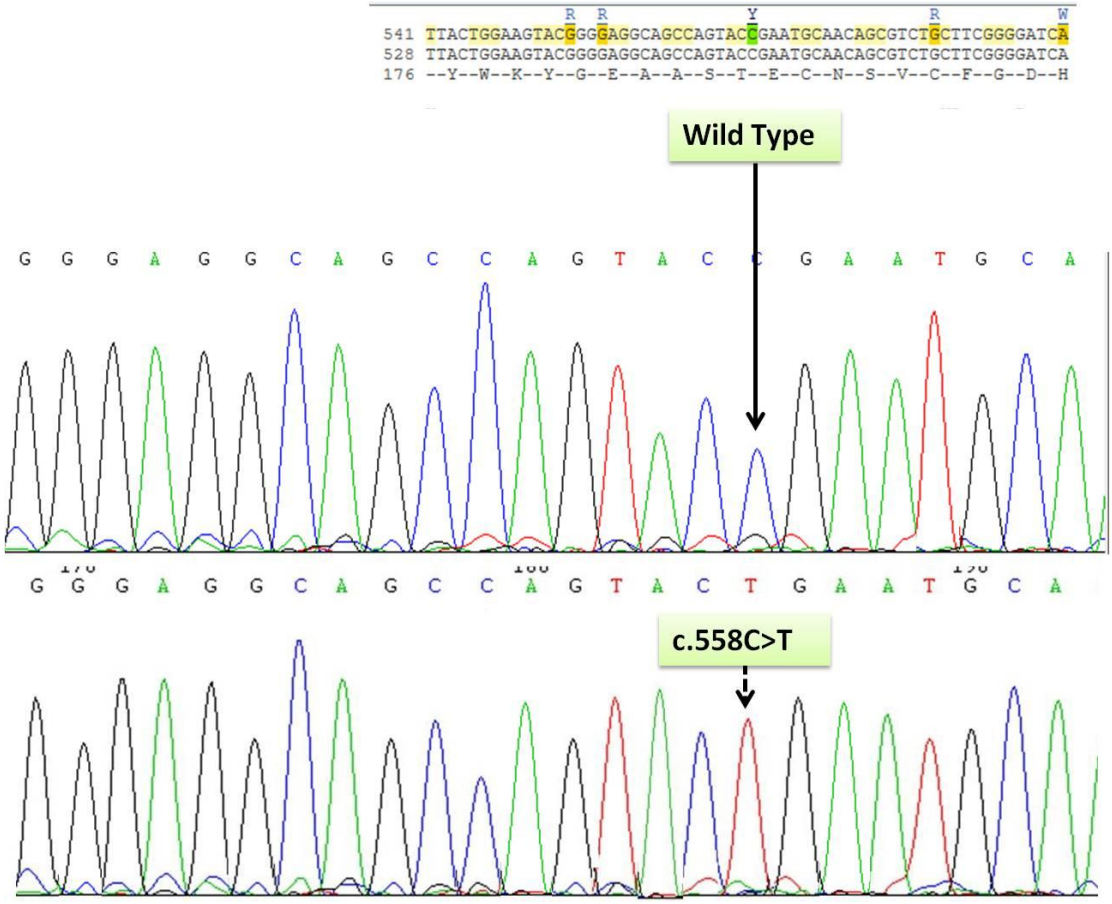


Figure 3.9: Polymorphism found in *KREMEN1*in affected member in *Family I*
Wild type compared to c.558C>T, which is a synonymous SNP p.T186T (TMP_ESP_22_29521331).

3.3.1.5 Microsatellite analysis

To follow up the shared homozygous regions on chromosome 19 and 22 microsatellites genotyping was performed on affected and unaffected members in *Family 1* for the candidate region (22467023-24169840) in chromosome 19, and for the candidate region 28200951-30560185 in chromosome 22 Figure 3.10. Primers used to amplify the microsatellite markers are listed in Appendix C2. The microsatellite genotyping analysis revealed that these regions

did not harbour the causative gene(s) as some of the unaffected were homozygous and some of the affected were heterozygous to the markers, as shown in Figure 3.11 and Figure 3.12.

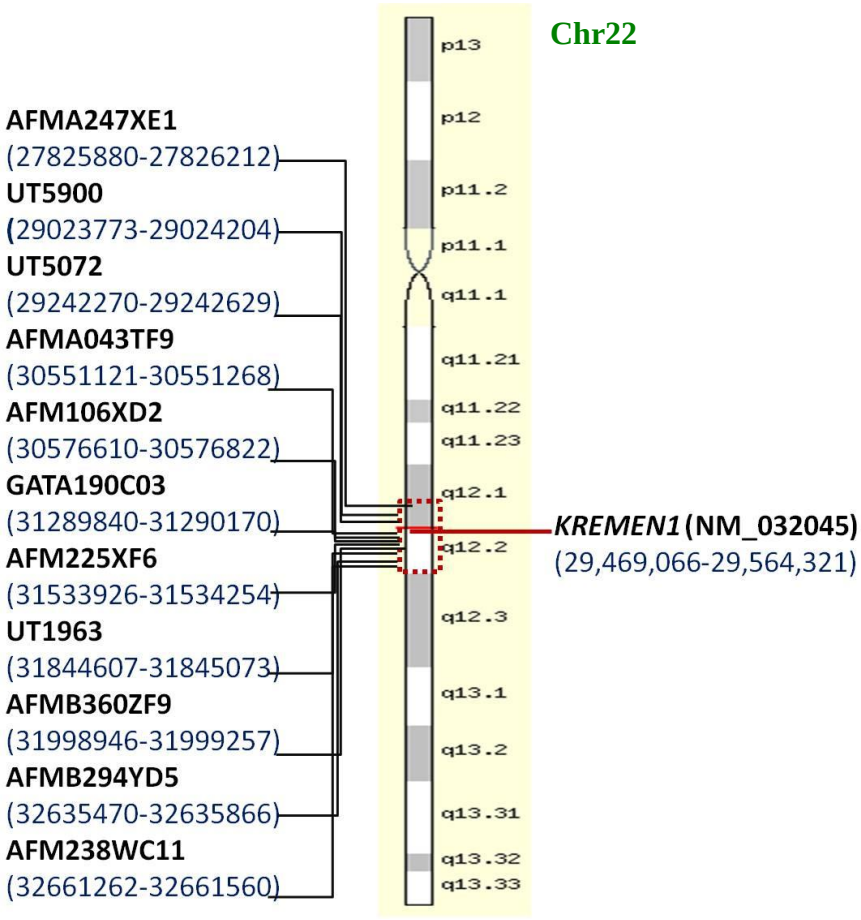
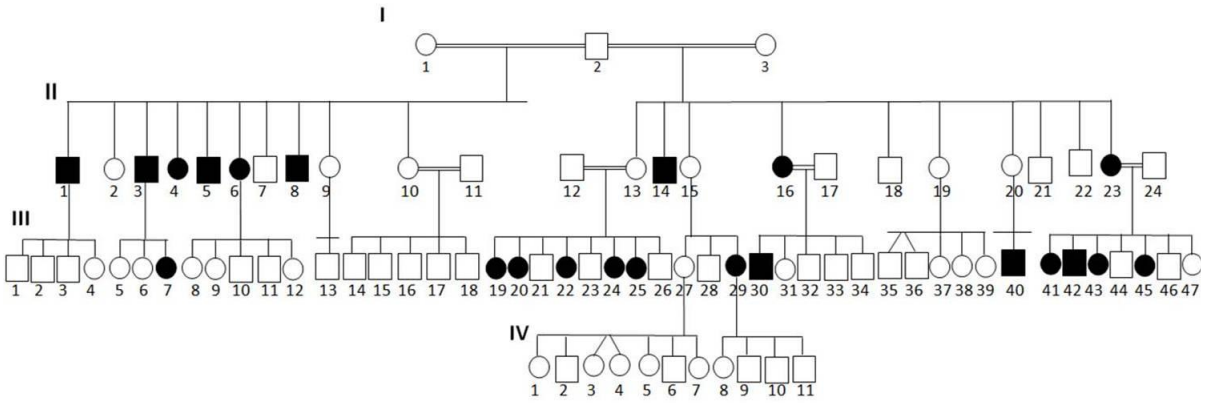


Figure 3.10: Chromosome 22 microsatellites map shows the position of *KREMEN1* gene in relation to the microsatellites markers



Marker	II:1	II:5	II:7	II:14	II:16	II:23	III:19	III:20	III:23	III:29
D19S1036	209/209	209/209	209/209	204/209	204/209	209/209	209/209	209/209	209/209	204/209
D19S434	271/283	271/271	271/271	267/271	271/283	271/271	271/271	271/271	271/271	271/283
S19S568	257/259	248/248	248/248	248/259	248/259	248/257	248/257	248/248	248/248	257/259

Figure 3.11 : Polymorphism markers in chromosome 19 for 10 members in *Family I*

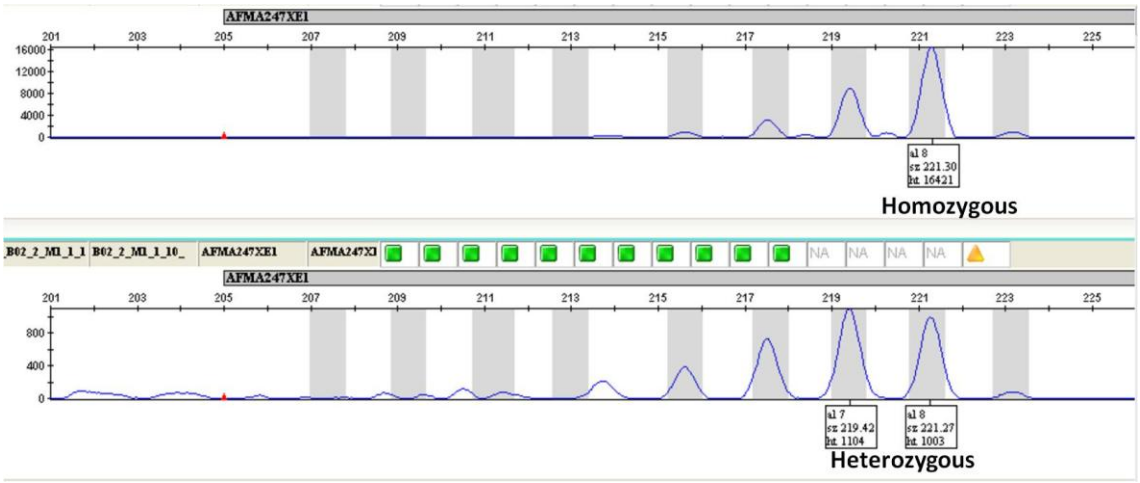


Figure 3.12: Gene Mapper® screen shots for the marker AFMA247XE1 in chromosome 22 showing an unaffected member who was homozygous and affected member who was heterozygous to this marker.

3.3.2 Discussion

3.3.2.1 Homozygosity mapping

Homozygosity mapping is a very powerful method for mapping genes which cause autosomal recessive disorders in consanguineous families (Kruglyak *et al.,* 1995). Many studies have revealed the power of homozygosity mapping as a first step in identifying the region of the genetic defect in recessive diseases such as congenital icthyosis in consanguineous population (Israeli *et al.,* 2013). This approach has also been used with non consanguineous populations to investigate the regions of known genes with diseased mutations such as retinitis pigmentosa (Collin *et al.*, 2011).

In this approach it is expected that all the affected individuals will be homozygous by descent at the disease locus, which means that they will also be homozygous by descent at closely adjacent loci. Therefore, if the family is analysed with multiple polymorphic markers distributed across the genome, and a homozygous region is found to be shared by all the affected individuals but not an unaffected control from the same family, then it is more likely that the disease locus is present within the shared region (Young, 2005). In this study, homozygosity mapping was carried out in a large consanguineous family (*Family I*), In this family, homozygosity mapping of eight affected members revealed many shared homozygous regions but none were large enough (1 Mb or more) to exclude the others. The most promising regions were in chromosome 19 and 22. Thus, more consanguineous families with similar phenotype were genotyped to identify the most likely homozygous regions that might harbour the causative gene(s). The genotyping analysis of the families showed that Family *V* and *VI* shared a relatively small region in chromosome 22 with *Family I* which was 29,305,513-29,635,842. Further investigation was carried out of the region in chromosome 22 by searching the region for candidate gene(s). In this region, only one gene was selected (*KREMEN1*) based on its function and studied by direct sequencing, this gene plays

important role in the Wnt signaling, it encodes the dickkopf homolog 1 (*DKK1*) and cooperates with it to block wingless Wnt/beta-catenin signalling (Cselenyi and Lee, 2008).

The failure of this strategy could be due the heterogeneity of the disease, or because of technical problems as there were some areas with no code and the causative gene might be in one of them.

3.3.2.2 Microsatellite genotyping

Microsatellite genotyping was used to examine large numbers of *Family I* members to see if the alleles of specific markers were inherited together with tooth agenesis. This approach was used because it allows investigation of large number of samples in one time at reasonable cost.

As shown in Section 3.4.1.5 the microsatellite genotyping of *Family I* unaffected members for regions in chromosome 19 and 22 revealed that some of them were homozygous and shared the same regions with the affected members, and some of the affected members were heterozygous for the investigated markers (Figures 3.11 and 3.12). As this approach failed to identify the causative gene(s), systematic protocol was used with these families based on the phenotype, the mode of inheritance, and the number of members available for genetic screening. However the simplicity of the phenotype of tooth agenesis in these families did not reflect the degree of its genetic complexity. Despite the broad application of homozygosity mapping and its great success in identifying the causative mutations in many autosomal recessive diseases, there are some conditions with simple or mild phenotype where homozygosity mapping might fail, as in the case of the families in this study.

3.3.3 Summary

In this study, homozygosity mapping of *Family I* did not reveal any large shared region. The members were homozygous in many small regions. The two most promising regions were investigated but none was found to be the right region. This suggested that this approach had failed and another strategy were needed to be taken to explore the molecular basis of *Family I*. Exome sequencing seemed to be the best choice as detailed clinical assessment of the family revealed that the phenotype and genetic basis might be complex. *Family I* was selected with two other families (*II* and *III*) in order to cover the general spectrum of tooth agenesis from mild to severe types.

3.4 Whole exome sequencing

3.4.1 Background

In this study, families *I, II* and *III* (batch 1) were selected for whole exome sequencing based on their diverse phenotypes. From each family, two affected members (Figure 3.12) were exome sequenced using a SureSelect Human all Exon 50 Mb kit (Agilent Technologies, Inc., Santa Clara, CA) and sequenced on an Illumina HiSeq to identify the causal gene(s). The percentage of exome captured in this experiment was between 87- 94% and revealed about 22203 variants or more in each sample. The sequence statistics are summarised in Table 3.4.

Candidate variants were filtered on the basis of function (as predicted by ANNOVAR), 1000 Genomes (www.1000genomes.org) and NHLBI exome sequencing projects (http://evs.gs.washington.edu/EVS/) frequencies. The candidate variants were selected to be either loss-of-function, non-synonymous or affecting a splice site junction, and a 0.5% frequency threshold was used for both 1000 Genomes and NHLBI.

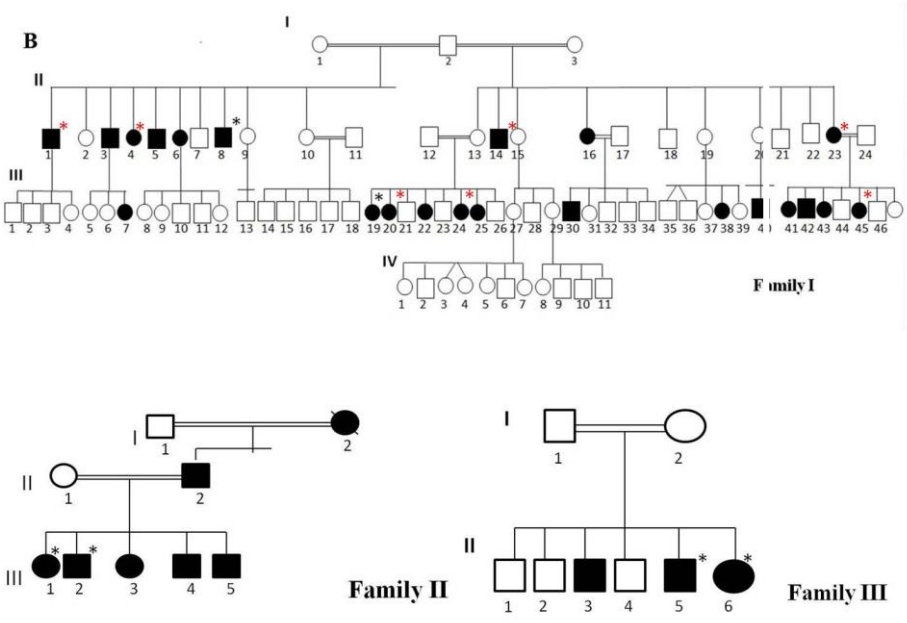


Figure 3.13: Pedigrees of the three families showing the sequenced family members (*) in batch 1(*) and batch 2

TABLE 3.4: Batch 1, Next Generation Sequencing statistics

Family's	***Family I***		***Family II***		***Family III***	
Member	**II:8**	**III:19**	**III:1**	**III:2**	**II:5**	**II:6**
Target size (pb)	33989968	33989968	33989968	33989968	33989968	33989968
Number of reads	58106733	54686045	48249504	52464105	60094288	46954773
% of exome captured with depth at least 2X	94.3%	94.1%	94.1%	94.1%	94.1%	93.5%
% of exome captured with depth at least 10X	88.6%	88.2%	87.1%	88.0%	88.3%	84.3%
% of exome captured with depth at least 20X	73.5%	71.2%	69.6%	74.8%	73.1%	60.6%
Mean target coverage	37.1X	35.6X	32.8X	37.3	37.7X	29.3X
Total variants	22966	22686	23468	23801	22203	22396

An additional seven affected members in *Family I* were selected for exome sequencing, two males (II: 1 and II: 14) and five females (II: 4, II: 23, III: 20, III: 24 and III: 45), from different branches in the family (Figure 3.13). Good quality DNA samples (Table 3.5) were exome sequenced using a SureSelect Human all Exon 50 Mb kit (Agilent Technologies, Inc., Santa Clara, CA) and sequenced on an Illumina HiSeq to identify the causal gene(s).

Table 3.5: DNA concentration of the samples quantified by Nanodrop ND-1000.

Sample ID	Nucleic Acid Conc.	Unit	A260	A280	260/280	260/230	Factor
II:1	358.6	ng/µl	7.173	3.885	1.85	1.11	50
II:4	170.7	ng/µl	3.414	1.89	1.81	1.14	50
II:14	1058.5	ng/µl	21.17	11.363	1.86	1.58	50
II:23	600.1	ng/µl	12.002	6.979	1.72*	1.03	50
III:20	350.6	ng/µl	7.011	3.808	1.84	1.36	50
III:24	211.7	ng/µl	4.235	2.19	1.93	1.26	50
III:45	716	ng/µl	14.321	8.06	1.78*	1.04	50

* sub optimal 260/280 (< 1.8)

3.4.2 Results

3.4.2.1 Family I (batch 1)

Family I presented with lower second premolar and upper lateral incisor tooth agenesis (Section 2.3.3.1). The analysis of the exome sequencing data for the two affected members, male (II: 8) and female (III: 19), presenting with the same phenotype (lower second premolar agenesis), were interrogated (for novel homozygous or heterozygous missense or loss of function shared variants), and revealed 56 rare variants (Appendix E). The candidate genes in this family were selected based on their reported function and type. Three variants shared between the two sequenced members were investigated in genes that function could be regarded as plausible candidate. One was a rare/novel variant in transcription factor *STRN3* (Striatin, calmodulin binding protein 3) in chromosome 6 which was not present in online databases dbSNP,1000 Genomes or NHLBI exome variant project (http://evs.gs.washington.edu/EVS/) (Figure 3.14).

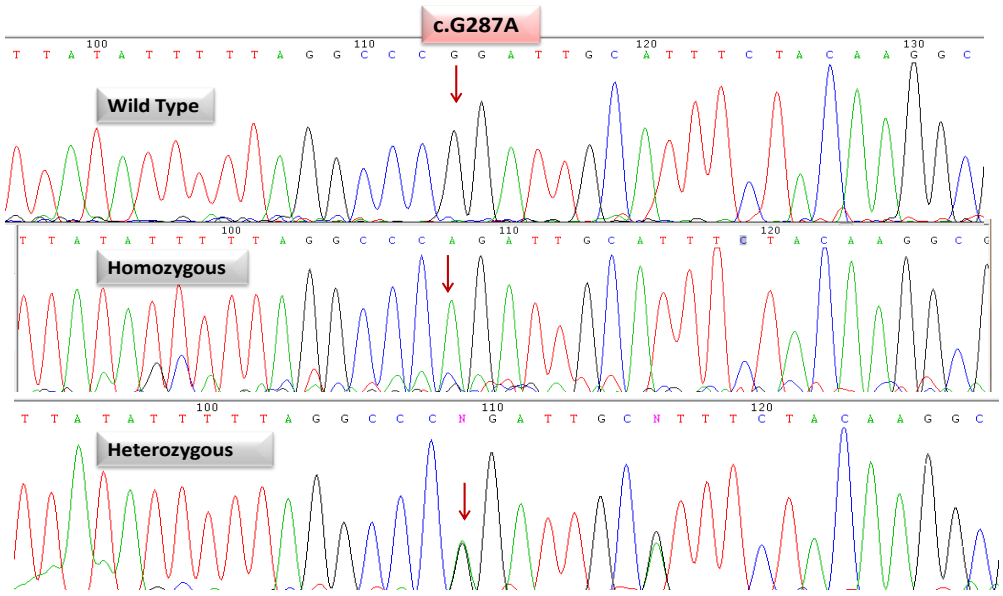


Figure 3.14: *STRN3* sequence electropherograms showing the homozygous (NM_014574.3:c. 287G>A.p.R96Q) mutation in exon 2 of the STRN3 gene in an affected member II: 8 compared to affected members with heterozygous (III: 19) and wild type (II: 5) mutation sequences.

The second rare variant was NM_015346.3:c.T2825C:p.M942T in *ZFYVE26* (Zinc finger, FYVE domain containing 26 OMIM 612012) in chromosome 14. The third rare variant was NM_004473.3:c.532_537del:p.178_179del in *FOXE1* gene (Forkhead box E1 'thyroid transcription factor 2' OMIM 602617) in chromosome 9, which was reported in dbSNP130 (rs71369530). These rare variants were sequenced as an interesting mutations shared among sequenced affected members in *Family I*. Genetic screening of other members revealed that this mutation did not segregate with tooth agenesis in the family (Figure 3.15).

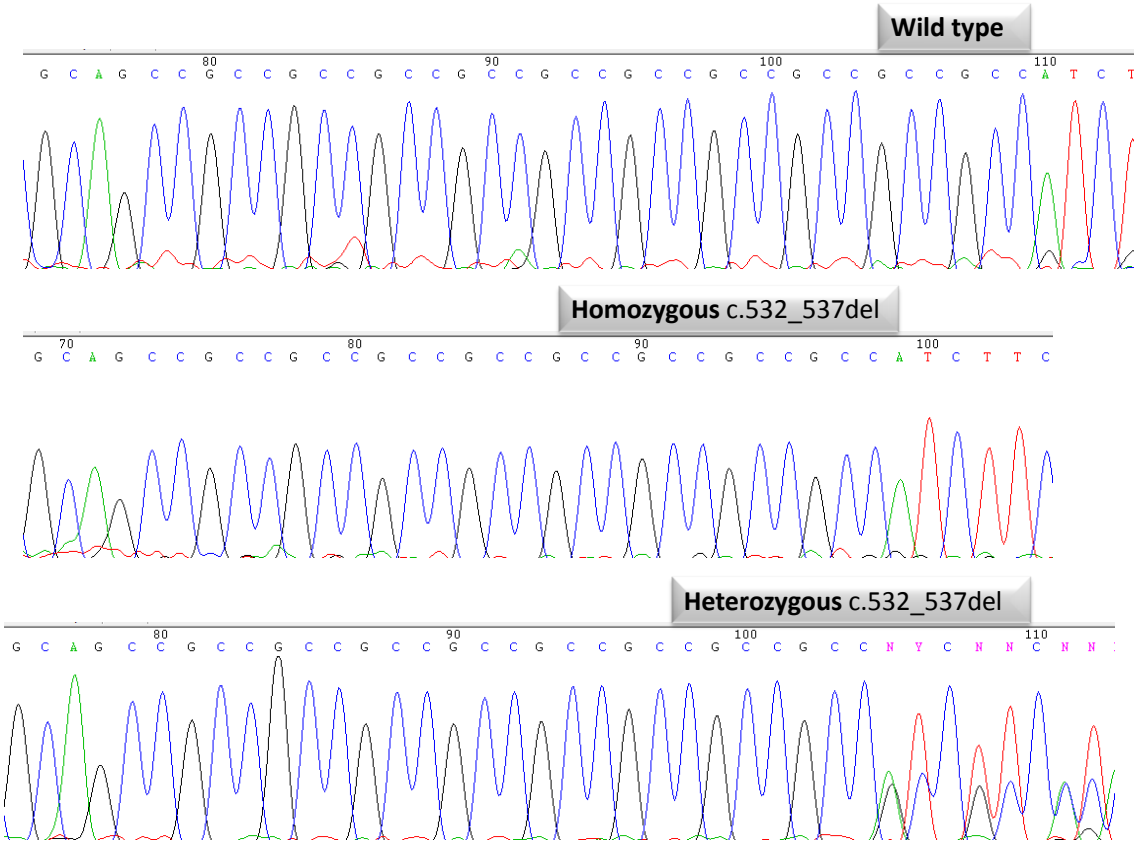


Figure 3.15: *FOXE1* sequence electropherograms showing the homozygous (NM_004473.3 c.532_537del:p.178_179del) mutation in exon 2 of the FOXE1 gene in an affected family member compared to affected family members heterozygous and wild type.

3.4.2.2 Family I (batch 2)

The percentage of exome captured in this experiment was much less than the previous, it was about 59% with depth at 2 and about 50% at 10 which revealed about 20690 variants or more in each sample. The sequence statistics are summarised in Table 3.6. The whole exome sequencing data in batch 2 did not reveal any shared rare variants between the seven members in *Family 1*.

Table 3.6: Batch 2, Next Generation Sequencing statistics.

Family's Member	***Family I***						
	II:1	**II:4**	**II:14**	**II:23**	**III:20**	**III:24**	**III:45**
Target size (pb)	62286318	62286318	62286318	62286318	62286318	62286318	62286318
Number of reads	66517910	67902072	77313924	89087678	69051580	73054382	91951110
% of exome captured (with depth at least 2X)	58.3%	60.5%	59.9%	59.2%	58.6%	58%	59.9%
% of exome captured (with depth at least 10X)	45.3%	52.7%	51.8%	51%	46.5%	43.3%	53.6%
% of exome captured (with depth at least 20X)	30%	43.6%	42.45%	41.45	32.%	30.9%	47.1%
Mean target coverage	15.795415	28.310824	27.16051	25.914468	17.6804	16.670703	37.432451
Total variants	20722	22067	22280	22122	21079	20690	22809

3.4.2.2.1 ***Homozygosity mapping using exome sequencing***

Homozygosity mapping was performed on nine samples from *Family I* (batch 1 and 2). The aim was to focus on the shared loss of function and non-synonymous variants in the shared regions that were captured by the exome sequencing. Most of the regions of homozygosity that were found by SNP genotyping were also observed by exome sequencing, although there was variability in the sizes of these regions. The obtained data did not show any shared rare variants, however a homozygous region in chromosome 17 was found shared by eight samples of the affected members (Figure 3.16). The minimum shared region between the

eight affected members was between 39580559-39670912. This region harbours a group of keratin genes including *KRT13*, *32*, *35*, *36*, *37* and *38*.

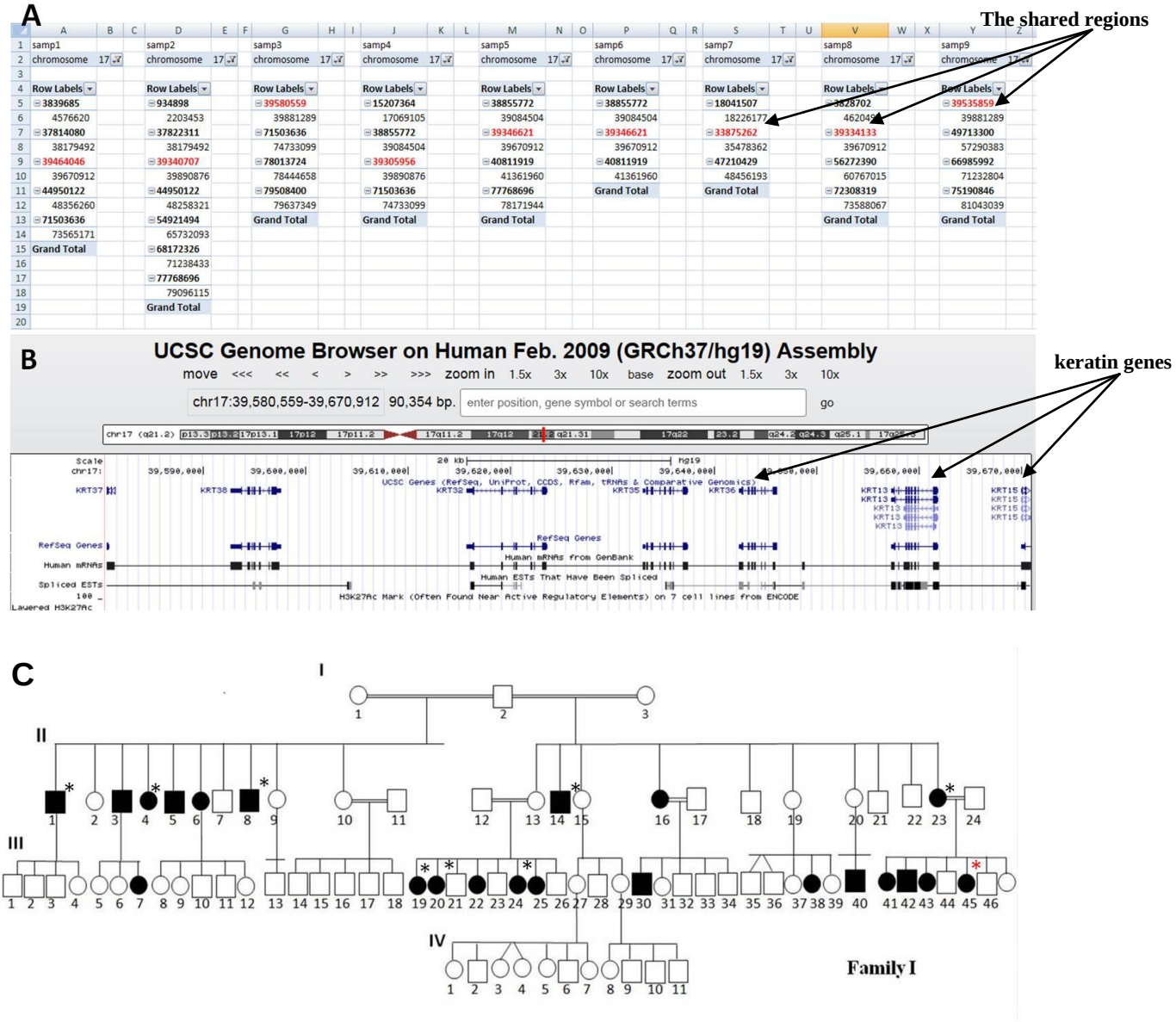


Figure 3.16: (A) shows the shared regions (colour red) between the samples in chromosome 17 in Excel sheet which were identified first. (B) UCSC Genome Browser shows the keratin genes in the minimum shared region 39580559-39670912 in chromosome 17. (C) Pedigree of *Family I* showing the affected members sharing the same region (*) and the one who was heterozygous to it (*)

3.4.2.3 Family II

Family II presented with autosomal dominant premolar and third molar agenesis (Section 2.3.3.2). Interrogating the exome sequence data of the two affected individuals (III: 1 and III: 2) in *Family II*, for novel heterozygous missense or loss of function variants shared by the two members, revealed that the two individuals shared 126 rare/novel variants, and that they were heterozygous for a 10 base pair (bp) frame shift insertion (c.750_751insACCGGCTGCC, p.F251PfsX92) in exon 2 in the *MSX1* gene (Muscle segment homeobox, drosophila, homolog of, 1, OMIM 142983) in chromosome 4.

These mutations are not present in the online databases dbSNP, 1000 Genomes or NHLBI exome variant project (http://evs.gs.washington.edu/EVS/). PCR and Sanger sequencing with the primer pair MAX1-F (ACTTGGCGGCACTCAATATC) and MSX1-R (CAGGGAGCAAAGAGGTGAAA) was used to confirm the segregation of c.750_751insACCGGCTGCC, p.F251PfsX92 with the dental disorder in the family.

Genetic screening of all the family members revealed that only the frame shift insertion mutation (NM_010835.2:c.750_751insACCGGCTGCC, p.F251PfsX92) segregated with the tooth agenesis in this family. This insertion was predicted to result in a frame shift from codon 250 leading to introduction of a new protein sequence that extended to the intron at codon 292. Thus the frame shift mutation is most likely the cause of the tooth agenesis in this family even if it was not in the DNA binding part of the MSX (Figure 3:17).

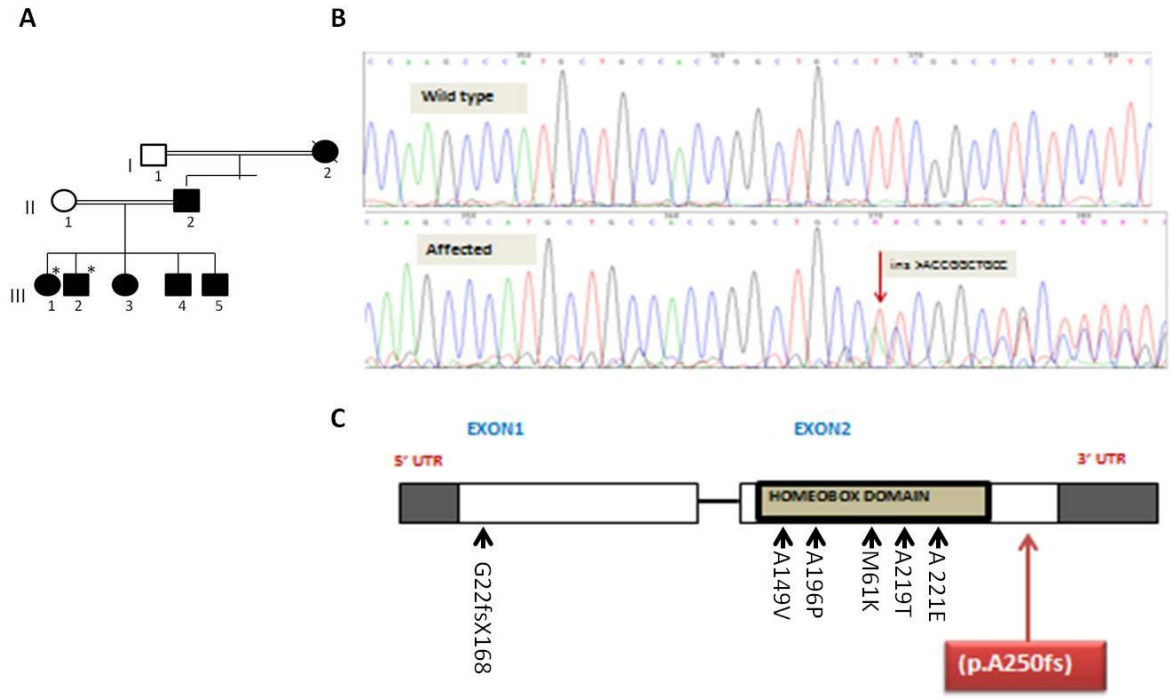


Figure 3.17: Genetic analysis of *Family II*.
(A) Pedigree of the family. (B) Sequence electropherograms of *Family II* showing the heterozygous (c.750_751insACCGGCTGCC, p.F251PfsX92) mutation in exon 2 in the MSX1 gene in an affected family member compared to control (wild type). (C) Schematic illustration of the structure of the MSX1 protein with homeobox domain and the approximate location of the identified frameshift mutation (red arrow). The black arrowheads point to the approximate positions of identified mutations in previous studies, A196P (Vastardis *et al.*, 1996), M61K (Lidral and Reising, 2002), A219T (Chishti *et al.*, 2006), G22fsX168 (Kim *et al.*, 2006), A149V (Mostowska *et al.*, 2006), and A221E (Xuan *et al.*, 2008).

3.4.2.4 Family III

Family III presented with autosomal recessive severe tooth agenesis (Section 2.3.3.3). The exome sequence data of the two affected members in the family (II: 5 and II: 6) were interrogated for novel homozygous missense or loss of function variants shared by the two affected individuals. This revealed a homozygous mutation, NM_022138.2: c.681T>A, in the *SMOC2* gene (SPARC related modular calcium binding 2 protein), changing a cysteine to a premature stop termination codon at codon 227 (p.C227X) in the mutant (Figure 3.18). This mutation is not present in the online databases dbSNP, 1000 Genomes or NHLBI exome

variant project (http://evs.gs.washington.edu/EVS/). PCR and Sanger sequencing with the primer pair SMOC2-F (GAGGACCACATCGTTCTTGG) and SMOC2-R (TGACTGTGTGTCGGTGTGTG) was used to confirm the segregation of c.681T>A (p.C227X) with the dental disorder in the family.

Genetic screening of all the family members revealed that p.C227X segregated with disease as a recessive mutation (Figure 3.17). These data clearly point to *SMOC2* being the disease-causing gene in this family. In addition, 44 rare/novel variants (Appendix E) were also identified as shared between the two affected individuals. For example, a homozygous novel frame shift deletion at c.818_912del (p.273_277del) which was identified in the Kalikrein10 (*KLK10*) gene was also investigated. However, the unaffected sibling (II: 4) was also homozygous for this *KLK10* mutation (Figure 3.19), excluding this variant as the underlying cause of hypodontia and microdontia. These data also suggest that loss of *KLK10* does not lead to an obvious developmental or clinical condition. None of the remaining rare/novel sequence variants were considered plausible candidates for the dental disorder.

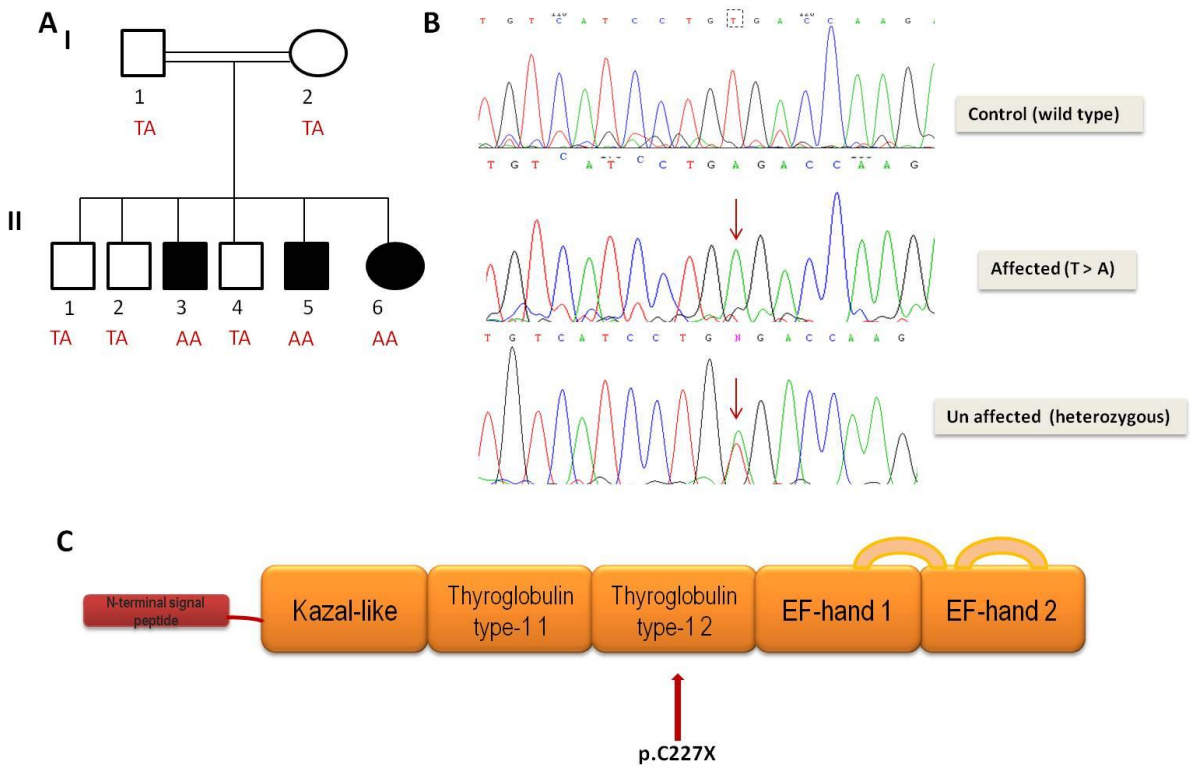


Figure 3.18: Genetic analysis of *Family III*.
(A) Pedigree of *Family III* showing the segregation of the mutation (NM_022138.2:c.681T>A:p.C227X) with severe tooth agenesis and microdontia. (B) Sequence electropherograms showing the homozygous c.681T>A:p.C227X mutation in exon 8 of the SMOC2 gene in an affected family member compared to unaffected family member (carrier) and unrelated control (wild type). (C) Schematic illustration of the structure of the SMOC2 protein with key function domains and the approximate location of the stop codon mutation leading to a truncated protein missing both EF-hand calcium binding domains.

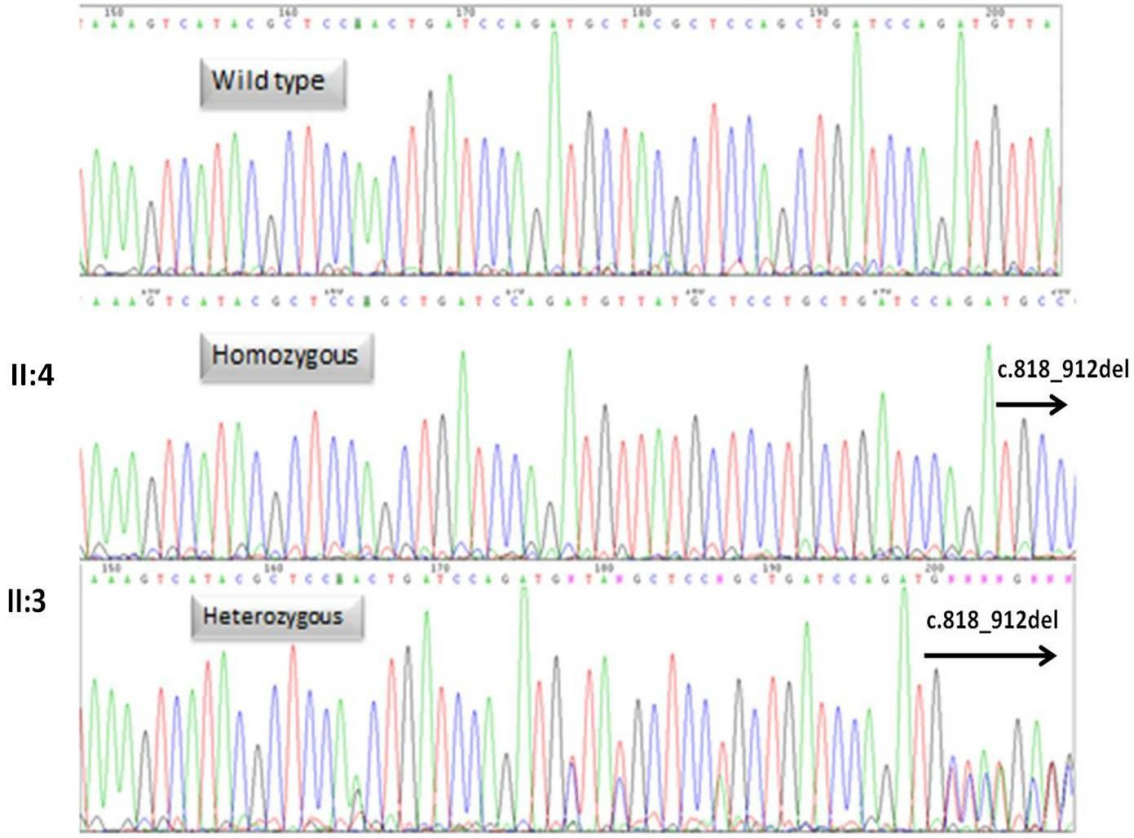


Figure 3.19: Sequence electropherograms of *Family III*, showing the homozygous NM_001077500.1:c.818_912del.p.273_277del mutation in exon 6 of the *KLK10* gene in an unaffected family member (II: 4) compared to a heterozygous affected family member (II: 3) and wild type.

3.4.3 Discussion

The ability to sequence the whole coding exome quickly, efficiently and cheaply has transformed human disease gene discovery and diagnosis, but the complete capture of all the exons by the sequence capture methods available is not 100%, so there will be some coding missing regions (Ku *et al*., 2012). In this study, exome sequencing was performed on three consanguineous families. These families were enrolled based on their phenotype. Family I presented with the most common mild phenotype of tooth agenesis, the family was missing mainly the lower second premolars (only two of the affected members having either upper lateral incisor tooth agenesis or malformed shape (peg shaped)) and segregated the tooth agenesis in an autosomal recessive manner. *Family II* presented with a specific phenotype of premolar and third molar agenesis in an autosomal dominant pattern; this phenotype has been reported several times recently with MSX1 mutations. The third family presented with clinical dental features similar to two previous cases reports, one being for a Pakistani family with severe tooth agenesis mapped to chromosome 16 (Ahmad *et al.,* 1998) and the other being a Turkish family with a comparable phenotype linked to SMOC2 in chromosome 6 (Bloch-Zupan *et al.*, 2011).

3.4.3.1 Family I

3.4.3.1.1 ***Batch 1***

In *Family I,* exome sequencing was performed on two affected quite distant relatives in which a recessive form of only bilateral lower second premolar tooth agenesis was observed. The percentage of exome captured in this experiment was about 94.2% with depth at 2X and 88.4 % with depth at 10X. Unexpectedly, exome sequencing did not reveal any of the shared rare/novel exonic homozygous variants, contrary to what might have been expected from any autosomal recessive condition in a highly consanguineous family of such a large size.

There were 56 rare variants shared among the two affected members (Appendix E). III: 19 was heterozygous in all of the rare variants whereas II: 8 was homozygous in a few of them (4 rare variants). From these variants, variants in STRN3 and FOXE1 genes were selected based on functional knowledge about them. For example transcription factor FOXE1 was initially identified in the thyroid and regulates the expression of thyroid-specific genes. Interestingly *Foxe1* mutant mouse was also found to have thyroid agenesis, spiky hair, and cleft palate. A role of FOXE1 has been proposed in mediating sonic hedgehog signalling Shh in the human epidermis of GLI2 (Eichberger *et al.*, 2004). Choi *et al.*, (2011) reported that MSX1 expression was downregulated contingent on FOXE1 downregulation in skin and keratinocytes (Choi *et al.*, 2011). These reported findings about FOXE1 were sufficiently convincing to justify the selection of the mutation in *FOXE1* as a candidate gene. But direct sequencing revealed that some affected members were wild type and did not have the mutation, *e.g.* II: 1 and II: 4 (Figure 3.15). This led to the exclusion of this variant as the causative mutation. The rare variant in *STRN3* was investigated mainly because it was shared between the sequenced affected members in *Family I* and *II*. Once more, the direct sequencing of other members in the families has excluded this variant from being the diseased mutation. For *Family I*, the study was first designed to search for a homozygous region that might harbour the diseased gene but this approach was less effective because of the complexity of the large family and the mild type of the phenotype. For this reason, whole exome sequencing was performed to see if any shared rare variants either homo or heterozygous.

3.4.3.1.2 ***Batch 2***

In batch 2, seven samples from different branches of the family were sequenced. The aim was to eliminate the number of shared rare variants between the previously two sequenced affected family members, but the results showed that the number of shared rare variants

across all sequenced individuals was zero. This unpredicted result in identifying the causative mutation in *Family I* may be either due to technical or analytical reasons or wrong genetic model.

In the second exome sequencing of *Family I*, the percentage of exome captured was relatively low, about 56% compared to the exome captured in batch I which was 94% (Table 3.6). The decrease in percentage of the captured exome can be blamed for the failure in identifying the causative rare variants, but the lack of shared homozygous rare variants between the two members in *Family I* in batch 1 which had a high percentage of exome captured should be considered as they shared 56 rare variants, 99% of which were heterozygous variants. Nevertheless, this method has some limitations. Homozygosity mapping using the exome sequencing data has revealed a shared small region among eight affected members; the 9th one (III: 45) had a homozygous region which was close to the shared one. This was a promising region especially since it was rich in keratin genes and the literature supports the hypothesis of keratin genes being responsible for tooth agenesis (Gass *et al*., 2009) (see Section 1.7.4). However the genotyping data of the III:45 member by *Illumina Infinium*® high-density array single nucleotide polymorphism (Section3.4.1.2) has excluded this region as it confirmed the heterozygosity of the III:45 member in this region. Interestingly the exome data of the nine members contained several mutations of keratin genes that were present in some members but missing in others. This finding might suggest that the incidence of simple tooth agenesis in this family might be part of a complex phenotype of a condition with other undetectable features that the researcher was not aware of. In complex phenotypes the clinical features are presumed to be an outcome of the accumulated effects of and interactions of multiple causative alleles, with a variety of genomic, environmental and local factors, as proposed in the multifactorial model (Brook, 1984). The effect sizes of the involved alleles in complex phenotypes are likely to vary in gradient ranges, from small and

indiscernible to large and significant, thus few alleles are expected to express a large effect that can be detected by the current approaches of molecular analysis (Marian, 2012b). Another approach should be considered to verify the molecular bases of tooth agenesis in *Family I*, either by whole genome sequencing or by a different strategy of linkage analysis (see Section 5.3).

3.4.3.2 Family II

Family II was first sequenced directly by Sanger regional sequencing for mutations in *PAX9* and *MSX1*. The results were negative although exome sequencing of the family revealed that the two sequenced affected members were heterozygous for a 10 base pair (bp) frame shift insertion (c.750_751insACCGGCTGCC, p.F251PfsX92) in exon 2 in the *MSX1* gene. This can be explained by the position of the mutation, which was at the end part of exon 2 after the highly conserved homeodomain sequence, and the DNA binding part (amino acids 167-225) (Figure 3.17). Also there is limitation of direct sequencing to identify the diseased mutation, because it depends on the designed primer for the gene and the overlapping of the amplified part to cover the whole gene exon, as well as the quality of the sequencing itself. The *Family II* phenotype was assessed from the beginning as a phenotype, consistent with previous studies that have linked it to mutations in the *MSX1*gene. The number of missing teeth was different among these families and even among *Family II* members, as the father was missing only four premolars and two third molars while his children had more missing teeth. This could be related to the haplo-insufficiency and allelic expression variability. However, this frame shift insertion does not seem to interfere with development of the lower first premolars as they were present in all affected members. All the affected members were missing the same tooth on both sides except one member (III: 2) who was missing only the upper left premolar. This led us to consider a key question about uni- and bi-lateral tooth agenesis when they exist in the same family and with the same mutation. This question was related to the

multiplicity of cell biological processes in the formation of the right and left body plan. Several genes have been identified that were asymmetrically expressed in the L/R orientation before LR asymmetric morphogenesis of the embryo. For example, genes such as *Nodal*, *Lefty2* and *Pitx2* are expressed in the left lateral plate mesoderm which is located on the left side of the embryo (Hamada *et al.,* 2002). Also *Pitx2* has been linked to the expression of the *Fgf8* (Tucker and Sharpe, 2004). In *Pitx2* knockout mice the expression of *Fgf8* is reduced and tooth formation is arrested at bud stage (Lin *et al.,* 1999; Lu *et al.,* 1999). Currently the leftward movement of nodal vesicular parcels has been suggested as responsible for the activation of the non-canonical hedgehog signalling pathway, in an asymmetric elevation in intracellular $Ca2^{+}$ and in changing gene expression (Hirokawa *et al.*, 2009).

Tooth development is a very complex process that involves hundreds of genes, but up to now, the role of most of them has not been identified. There are some genes that have been identified for their mandatory function in the initiation stage of tooth development in knockout mice studies (Thesleff and Sharpe, 1997; Zhang *et al.,* 2005; Cudney and Vieira, 2012). These are signals derived from the oral epithelium (FGF8 and BMP4) and transcription factors that include MSX and DLX families in the mesenchyme (Wang *et al.,* 2011). These studies showed that in more than one member of each family these are expressed in the developing teeth and they compensate each other. Therefore the inactivation of both *Msx1* and *Msx2* or both *Dlx1* and *Dlx2* in mice arrested tooth formation at the initiation stage (Thomas *et al.,* 1997; Satokata *et al.,* 2000). Wang *et al,*. (2011) investigated the assumption of synergism of MSX1 with PAX9 in tooth agenesis by studying five known *MSX1* missense mutations linked to tooth agenesis but could not prove any synergism in MSX1 with PAX9 in these mutations, which suggests the independency of *MSX1*gene in causing tooth agenesis. Also they found that none of the molecular mechanisms they used

yielded a satisfactory explanation for the pathogenic effects of the *MSX1* mutations and recommended looking for a different approach in investigating this step of odontogenesis at the molecular level (Wang *et al.,* 2011).

At present, 12 mutations in MSX1 gene have been identified with tooth agenesis. Three were in exon 1 (van den Boogaard *et al.,* 2000; Lidral and Reising 2002; Kim *et al.,* 2006), seven were in exon 2 (Vastardis *et al.,* 1996; Jumlongras *et al.,* 2001; De Muynck *et al.,* 2004; Chishti *et al.* 2006; Mostowska *et al.,* 2006b; Mostowska *et al.* 2012; Xuan *et al.*,2008), and two were gene deletions (Nieminen *et al.,* 2003) and intronic deletions (Pawlowska *et al.,* 2009). The reported mutations in MSX1 were generally inherited in an autosomal dominant pattern apart from the one described by Chishti *et al.,* (2006) where a missense mutation segregating in a consanguineous Pakistani family in a recessive pattern was found. All these mutations, syndromic or non-syndromic, caused severe tooth agenesis, with only one exception reported by Mostowska *et al.,* (2012), in which the heterozygous missense mutation at position 224 caused selective tooth agenesis that involved only second premolars and third molars. *Family II* was almost similar to the previous case reports with regard to the mode of inheritance and type of teeth involved.

There was an interesting observation was made for *Family II*. Although the frameshift mutation found in this family was not in the highly conserved homeodomain sequence (amino acids 167-225), it produced agenesis in the premolars and third molars, which means that this part of the *MSX1* gene may has a specific function for the premolars and molars that has not been reported before.

3.4.3.3 Family III

In *Family III*, exome sequencing was performed on two affected siblings in which a recessive form of severe tooth agenesis and microdontia was segregating. The percentage of exome captured in this experiment was about 93% with depth at 2X and 88 % with depth at 10X, which revealed 44 rare/novel exonic homozygous variants. The exome sequence analysis of the two affected siblings in the family revealed that they shared a homozygous premature stop codon mutation in exon 8 (c.681T>A, p.C227X) of the *SMOC2* gene encoding the SPARC related modular calcium binding 2 protein. SPARC related modular calcium binding 2 protein, OMIM 607223, SMOC2 is a member of the BM-40 protein family. A matricellular protein is a group of non- structural extracellular proteins that provide regulatory roles by modulating cell-matrix interactions and cell functions. These proteins are characterised by containing binding sites for extra cellular matrix (ECM) structural proteins and cell surface receptors, and seem to have a role in modulating activities of specific growth factors. Many of these proteins play a major role in wound healing and tissue repair such as the secreted protein acidic and rich in cysteines (SPARC) family members which includes SPARC related modular calcium binding 1 and 2 proteins (SMOC1 and SMOC2), (Vannahme *et al*., 2003).

SMOC2 was first isolated from bone but later found expressed in many other tissues. Experimental studies in mice showed that deletion of this BM-40 gene does not alter embryogenesis but developed cataracts, severe osteopenia and also displayed defects in wound healing and in tissue remodelling (Vannahme *et al*., 2003). It has been reported that SMOC2 is expressed in early embryogenesis and can interact with extracellular matrix (ECM) receptors such as αvβ1 and αvβ6 integrins (Liu *et al.,* 2008). In mouse embryo the expression of this gene has been found high in the craniofacial region at embryonic day 14.5

(E14.5) which is the begining of the cap stage of tooth development, mainly at the level of the tooth mesenchyme (Bloch-Zupan *et al.*, 2011).

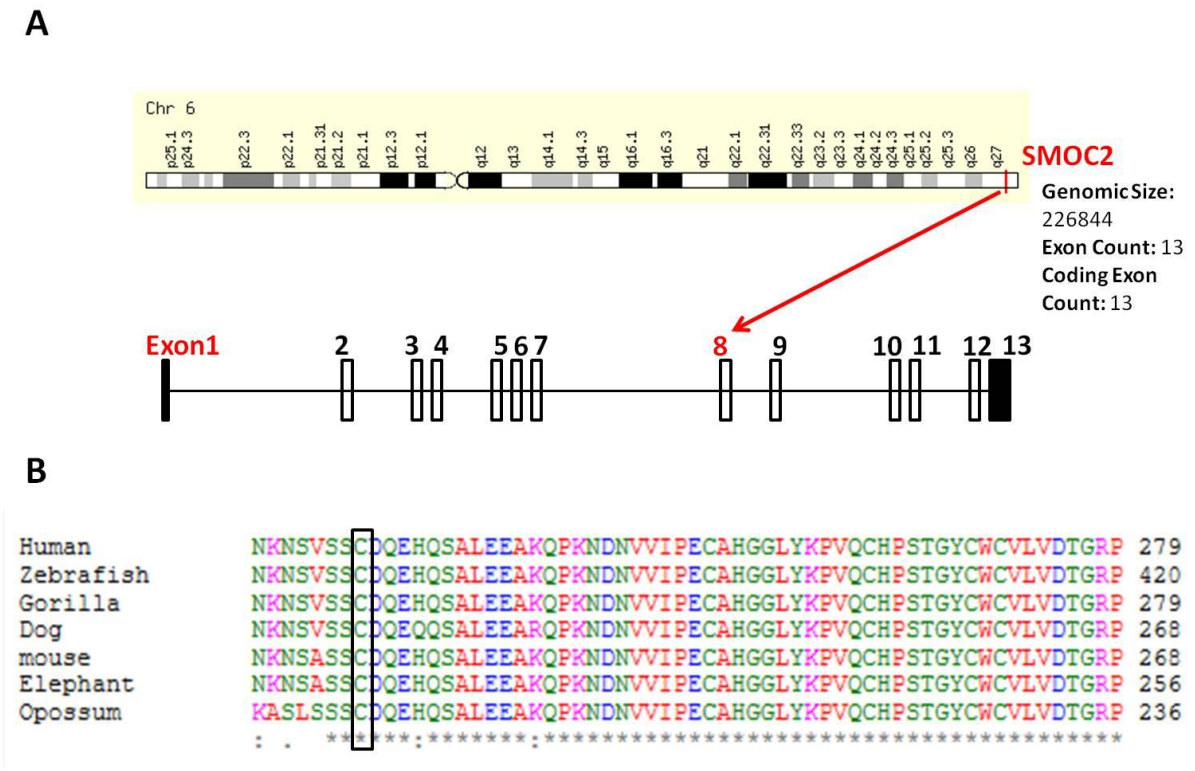


Figure 3.20: *SMOC2* gene
(A) Schematic illustration of exon-intron structure of *SMOC2* gene shows the exons with mutations (1 & 8). (B) ClustalW alignments of SMOC2 protein sequences of different species show the cysteine residue at codon 227.

Previously, a homozygous splice-site mutation (c.84+1G>T) in *SMOC2* gene localized at chromosome 6q27-ter, has been reported in a clinically similar family of Turkish ethnicity (Bloch-Zupan *et al.*, 2011). The SMOC2 gene has 13 exons and the stop codon mutation found at exon 8 (Figure 3.20) would lead to truncated protein, missing thyroglobulin type-1 2 and the calcium binding domains EF-hand 1and EF-hand 2.

Although many studies have been published about the role of genes in odontogenesis and about mutations in these genes, links with tooth agenesis and their specific role in tooth development is still an enigma. Therefore further studies identifying mutations in different

parts of such genes like the *MSX1* and *SMOC2* genes are needed. It is well established that the four signalling pathways, TGF, FGF, Hedgehog (SHH) and Wnt, are key regulators of tooth development. Any disturbance within these signalling pathways will potentially affect or block normal tooth development (Thesleff, 2003). Thus, although the function of SMOC2 in mammalian development has yet to be fully investigated, there is some evidence that it may affect the receptor-mediated signalling of many growth factors (Rocnik *et al.,* 2006; Liu *et al.,* 2008). Additionally, SMOC2 expression was dependent on Hedgehog signalling in foetal gonad and reproductive tract differentiation (Pazin and Albrecht, 2009).

Many researchers have reported that tooth agenesis is associated with other dental anomalies such as microdontia, which was the most commonly associated anomaly described (Baccetti, 1998, Tallon-Walton *et al.,* 2007; Brook *et al.,* 2009a), but few studies have investigated the genetic basis of the correlation between tooth shape in general or reduction in tooth size (microdontia) with tooth agenesis. Little information is available about the association between these dental anomalies. Interestingly, both Pakistani and Turkish families harbouring *SMOC2* mutations presented with similar dental anomalies: oligodontia, microdontia with very globular and malformed teeth and incisal notches, taurodontism of the molar teeth, and short roots. However, *Family III* did not express macrodontia as reported in the Turkish family (Bloch-Zupan *et al.,* 2011). As macrodontia is rarely associated with tooth agenesis, the clinical difference between the two families may be related to another gene mutation/variant, or due to the distinct *SMOC2* mutations segregating in each family. An interesting clinical feature in both families harbouring *SMOC2* mutations is that the deciduous teeth and permanent molars (excluding third molars) are present. The shapes and sizes of the permanent molars and the deciduous molars are affected similarly, indicating that

the development of permanent molars might be regulated by the same genes as the deciduous series.

The exome sequencing of the three families did not reveal any mutation in *WNT10A.* This finding is not consistent with what Van den Boogaard and his co-workers found. They identified mutations in *WNT10A* in about half of their tooth agenesis cases with other pathogenic mutations in MSX1, PAX9 and AXIN2 (Van den Boogaard *et al.,* 2012).

3.4.4 Summary

The molecular basis of tooth agenesis in *Family I* was investigated by homozygosity mapping and whole exome sequencing, the most advanced genetic approaches currently available in genetic research (Ritchie *et al.,* 2011, Ku *et al.,* 2012). The failure of those approaches in detecting the causative gene(s) in *Family I* might have been resulted from the complexity of the condition even though it appeared to be a simple mild phenotype. The same approach was successful in locating the genes in other families (*Family II* and *III*) which had a more severe phenotype with regard to the type and number of teeth missing. The results in this study show that severity of the phenotype increases the possibility of identifying the causative mutation. This suggestion is supported by the number of identified mutations in severe tooth agenesis cases compared to moderate or mild cases (Das *et al.,* 2003, Chishti *et al.,* 2006, Xuan *et al.,* 2008, Pawlowska *et al.,* 2009, Ayub *et al.,* 2010, Mostowska *et al.,* 2012). The *MSX1* frameshift mutation was linked to a family with moderate phenotype. This verified the development of the second premolar and third molar teeth is affected by the end part of the *MSX1* gene which was not in the highly conserved homeodomain sequence. The *SMOC2* homozygous premature stop codon mutation was found in a family with severe phenotype. This present study is the only other report on the role of the SMOC2 gene in tooth development apart from that reported by Bloch-Zupan *et al.* (2011).

CHAPTER 4: TOOTH CROWN DIMENSIONS IN *FAMILY I* (INITIAL STUDY)

4.1 Background

Many studies have reported an association between tooth agenesis and crown dimension reduction or microdontia (Garn and Lewis 1970; Baum and Cohen, 1971; Rune and Sarnas 1974; Brook, 1984; Schalkvanderweide *et al.* 1994), proposing that these condition might have similar aetiologies. However there are an increasing number of pieces of evidence and case reports to suggest a connection between tooth agenesis and the crown size of the remaining dentition. This assumption has been investigated thoroughly among individuals with different type of tooth agenesis and their relatives (McKeown *et al.,* 2002). Many authors have noticed this link mainly in severe types of tooth agenesis, based on the number of missing units (Brook *et al.*, 2009b ; Bloch-Zupan *et al.*, 2011). Few studies have attempted to investigate the correlation between reduction in tooth size and the specific missing tooth type.

To date, various methods have been developed to measure the dimensions of human teeth. Each of these methods has advantages and disadvantages, and the choice of which to use depends on the study design, aims, cost and time budgets. These methods can be classified into classical and advanced methods or direct and indirect methods.

The manual technique has been considered as the gold standard of tooth size measurement methods. This technique obtained linear measurements by the use of hand instrument such as Boley gauge and calliper and classified as a direct measurement technique (Moorrees and Reed, 1964, Garn and Lewis, 1970, Brook, 1974, Brook et al., 2009b). Image analysis system technique which was proposed by Brook in 1983(Brook *et al.*, 1983), has been compared to the manual method, and was shown to be more variable than the manual technique (Brook *et al.*, 1986).

The new technologies of imaging and analysis have been used lately to describe and define the tooth size in a non destructive method. There are several three-dimensional image analysis system available in the market and each system has its own processing software to create a full surface rendered 3D models (Smith *et al.*, 2009).

In order to improve our understanding of the aetiology of these observations, this study focused on only one large family (*Family I*, Figure 4.1 and 4.2), and studied the tooth dimensions of their members.

In this family, the assumption was that the tooth agenesis resulted from one mutated gene that segregated in the family members Figure 4.1. Other families in this study were not included for standardisation as the causative mutations in these families might be different, therefore their effects might be different. A further investigation using imaging three dimensional (3D) technique has been considered (see Section 5.3).

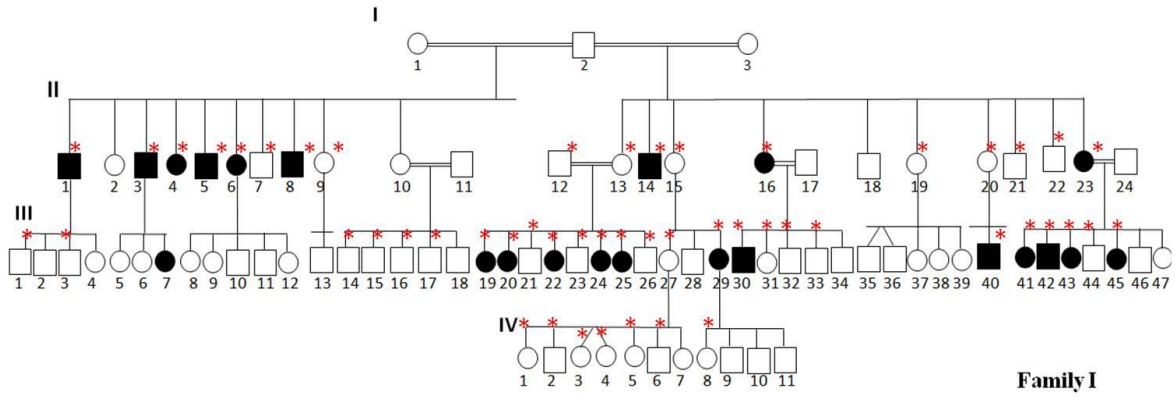


Figure 4.1: *Family I* pedigree showing the study group: 21 affected members (8 males and 13 females) and 29 relatives without tooth agenesis (17 males and 12 females).

Figure 4.2: Frontal clinical views of *Family I* members, males (II: 3, II: 8, II: 18, and II: 26) and females (II: 4, II: 16, II: 20, and III: 43) showing the alignment and size of some of the affected and their relatives.

4.2 Materials and methods

4.2.1 Sample size

Following ethical approval by the ethics committee in the National Guard, Health Affairs, and the QMUL research ethics committee, a study of *Family I* members affected with a selective tooth agenesis, their first degree relatives and a control group was carried out by measuring mesio-distal (MD) and bucco-lingual (BL) dimensions in all erupted teeth excluding the third molars. Fifty family members gave informed consent to the study providing 21 affected members (8 males and 13 females) and 29 relatives without tooth agenesis (17 males and 12 females). The control group consisted of 163 sets of dental casts of 82 males and 81 females from dental clinics in King Abdul Aziz Medical City, National

Guard, Riyadh, Saudi Arabia, with complete permanent dentition and no family history of tooth agenesis. A single operator (the researcher), according to manufacturer's instructions, took upper and lower impressions of vinyl polysiloxane impression materials (Express ™ 2 Penta ™ Putty Soft, 3M ESPE AG, dental products, D-82229 Seefled, Germany) using Pentamix ™ 3 mixing unit (3M ESPE AG, dental products, D-82229 Seefled, Germany) and cast in blue dental stone. The MD and BL dimensions of each tooth were measured twice on different occasions.

4.2.2 Selection criteria

For the control group, all individuals had to be of Saudi ethnicity and the age group was 13-55 years. Carious or restored teeth, teeth with defects on the study models and teeth which were not completely erupted, deciduous teeth and third molars were all excluded from the study. None of the affected members in *Family I* had any associated syndrome or had unexplained previous loss of teeth. All family members completed medical questionnaires.

4.2.3 Crown measurements

Hand measurements of mesio-distal (MD) and bucco-lingual (BL) dimensions were obtained from study models of the 213 individuals using a digital caliper (Professional Caliper, Wenling Lifeng Precision Tool Co. Ltd., Zhejiang, China). The accuracy of measurement technique was 0.01mm (Figure 4.3).

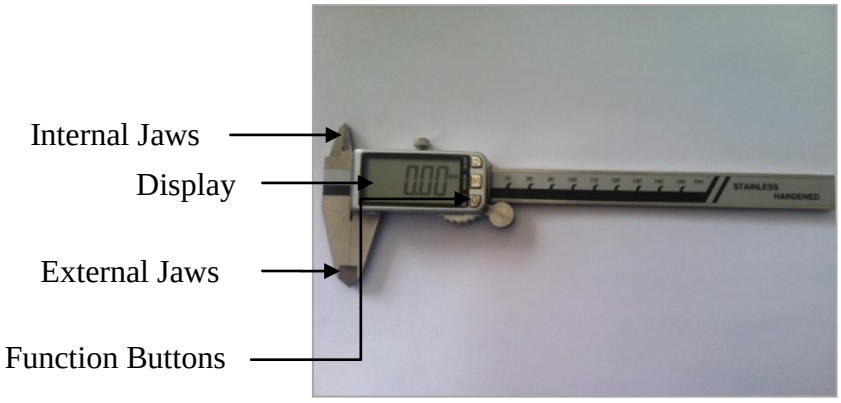


Figure 4.3: A professional caliper, Wenling Lifeng Precision Tool Co. Ltd., Zhejiang, China

Measurements were performed methodically from the upper left quadrant to the upper right quadrant, and from the lower left quadrant to the lower right quadrant according to the method described by Brook, Griffin *et al.*, (2009a). Each tooth was measured twice, at different times, and the mean value of the two measurements was used.

4.2.4 Mesio-distal (MD) measurement

The mesio-distal (MD) distance was measured as the maximum distance between the contact points on the proximal surface of the tooth crown, as defined by Moorrees and Reed (1964), with the caliper held parallel to the occlusal and buccal surface (Figure 4.4.A).

4.2.5 Bucco-lingual (BL) measurement

The bucco-lingual (BL) distance was measured as the maximum distance between the labial and lingual surfaces of the tooth crown in a plane perpendicular to that MD measurement (Figure 4.4.B).

4.2.6 Repeatability and reproducibility

Intra-operator repeatability was determined by measuring the MD and BL dimensions of 10 upper and lower casts twice at a different time.

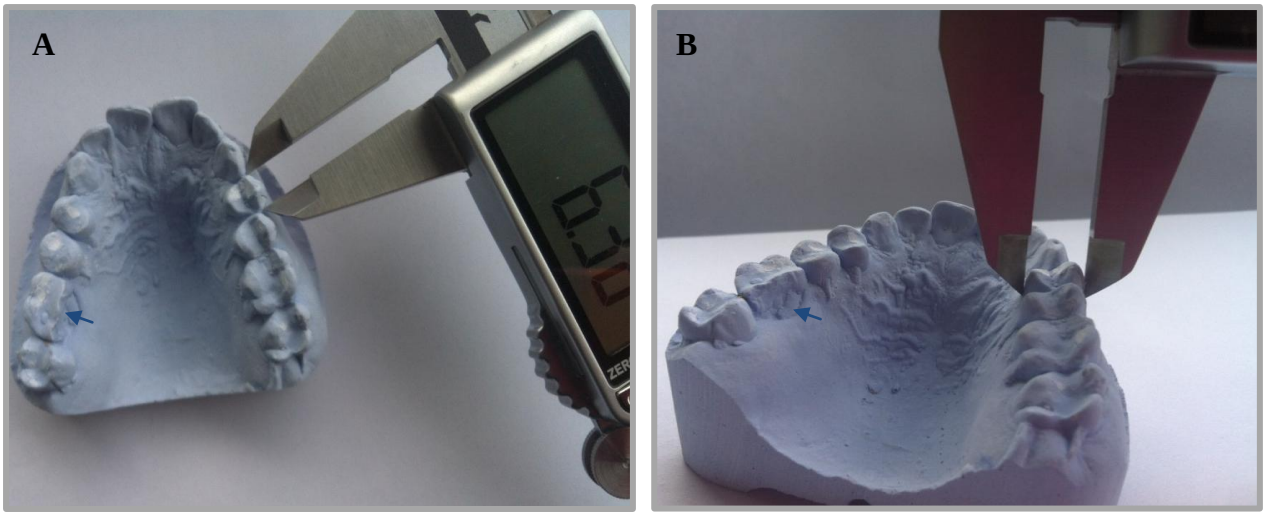


Figure 4.4: **MD and BL measurements.**
(A) MD distance (the maximum distance between the contact points on the proximal surface of the tooth crown). (B) BL distance (the maximum distance between the labial and lingual surfaces of the tooth crown in a plane perpendicular to that MD measure.

4.2.7 Statistical analysis

Statistical testing was performed using SPSS® software version 16. Reliability was assessed statistically using paired t-tests. The mean value of two measurements of the study groups was used. The comparison between the affected groups and their relatives without tooth agenesis might be questioned as the affected individuals could not be regarded as independent due to shared family membership. However the main value of the three groups (affected, the relatives and control) was compared using simple t-tests of the affected and related vs. controls, and also using the ANOVA to compare the three groups as independent groups. For the purpose of this study the level of significance was set at p value = 0.05.

4.2.7.1 Paired t-test (two-tailed)

To assess the reliability of the methodology in this part of the study, a two-tailed paired (dependent samples) t-test was run in SPSS v.16, to measure the difference between the first and second measurements (intra and inter operator repeatability and reproducibility).

To test the sample normal distribution as required for the paired t-test, a Levene test for equal variance was used. The t-test has very strong statistical power and it is very sensitive to differences between sets of measurements, as it tests whether the sample means of two sets of measurements are significantly different, with the two sets of measurements considered in a pair-wise fashion.

4.2.7.2 Independent samples t-test

In order to assess the possible effect of tooth agenesis mutation on the MD and BL tooth crown dimensions, participants were categorised in two groups (the affected and their relatives and the control group). An independent samples t-test was used to compare the mean values of each tooth between the two groups. To test the sample normal distribution, a Levene test for equal variance was applied.

4.2.7.3 Analysis of Variance (ANOVA)

A one-way analysis of variance (ANOVA) was used to assess the influence of the tooth agenesis mutation on the MD and PL tooth crown dimensions. This test is used where the means of more than two groups need to be compared; in this study three groups were examined. The Levene test for equality was performed first to test the normal distribution of the sample. ANOVA produces a test statistic (F) which assesses whether the group means are equal. A significant F value indicates that there is a significant difference between the groups, but it does not distinguish exactly where the difference lies. To identify which specific groups were significantly different, post hoc tests were performed within the one way ANOVA, for the significant values.

4.3 Results

4.3.1 Intra-operator reliability

In order to validate the accuracy of the measurement method, five upper and five lower (10) dental casts were selected at random from the study group. The overall Fleiss intra class correlation coefficient and R coefficient were > 0.95 and > 0.963 respectively. Paired t-tests revealed no statistical difference between the two sets of measurements (p-value >0.05 for all teeth measurements). Summary statistics are shown in Appendix F. The test results show that intra-operator reliability was excellent.

The normality test by male and female for each tooth within each group showed that the frequencies followed a normal distribution pattern. In all teeth the Leven test for equality of variances had a P value that exceeded 0.05. Moreover, the Kolmogorov Smirnov KS test P value was greater than 0.05 for most of the teeth measurements. This confirmed that variances between the groups were equal and satisfied conditions for the ANOVA and independent t-test.

4.3.2 Comparison of tooth crown measurement variables between the study group (affected and unaffected) and the control group using the independent t- test

Analysis by independent t-test revealed that in the male group, 26 teeth out of 28 teeth were significantly smaller in the study group ($p \leq 0.05$) in the MD and/or BL widths. In the female group, only 15 teeth were significantly smaller in the study group ($p \leq 0.05$) in the MD and/or BL widths. The means of the male study group were significantly smaller than the control group for all variables except the MD in tooth No. 13, 23, 25, 32, 33, 35, 41, 42 and 45, and the BL in tooth No. 17, 26, 31, 32, 33, 43, and 46. For the females, the means of the MD widths of the study group were significantly smaller than the control group for tooth No. 14, 12, 24, 25, 27, 34, 36, 37, 43, 44 and 47, and the means of the BL measurement of the study

group were significantly smaller than the control group for tooth No. 14, 12, 11, 21, 22, 25, 34, 35, 36, 37, and 47 (Figure 4.5 and Table 4.1 (1, 2, 3 and 4).

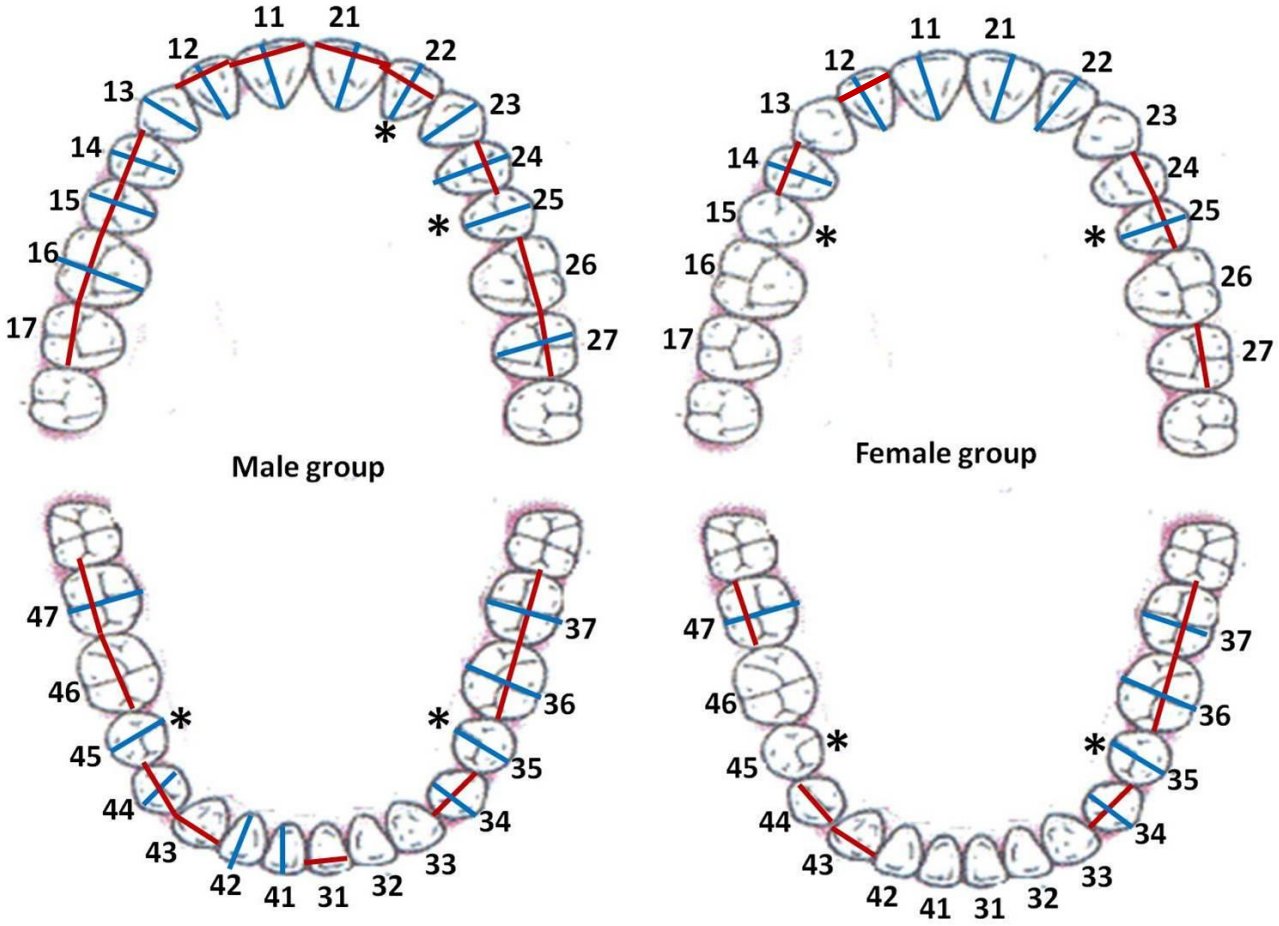


Figure 4.5: The lines indicate that the dimensions of the tooth types are significantly ($p \leq 0.05$) smaller in the study group for male and female based on independent t-test analysis (red line indicates MD dimension, blue line indicates BL dimension, * indicates congenital missing teeth in the affected group).

Table 4.1.1: Independent t-test for MD and BL means of upper right quadrant in male and female. Significant differences in group means are shown in red.

Statistics of upper right quadrant									
		Male				**Female**			
	Study group	**n**	**Mean**	**Std. Deviation**	**p-value**	**n**	**Mean**	**Std. Deviation**	**P-value**
T17_M_D_mean	Study	15	9.8757	0.61314	**0.004**	17	9.2985	0.50472	0.063
	Control	65	10.4518	0.69825		59	9.5936	0.58535	
T17_B_L_mean	Study	15	11.257	0.54387	0.542	17	10.5012	0.70331	0.156
	Control	65	11.3853	0.76682		59	10.79	0.74014	
T16_M_D_mean	Study	25	10.0932	0.5134	**0.000**	24	9.8808	0.57771	0.367
	Control	76	10.6376	0.63152		72	10.0108	0.61726	
T16_B_L_Mean	Study	25	11.3438	0.63782	**0.033**	24	11.089	0.60797	0.237
	Control	76	11.6661	0.64965		72	11.2562	0.59315	
T15_M_D_mean	Study	24	6.3865	0.36843	**0.000**	17	6.3144	0.4801	0.061
	Control	77	6.8923	0.50749		73	6.5632	0.48816	
T15_B_L_Mean	Study	24	8.9883	0.41478	**0.000**	17	8.9121	0.68172	0.117
	Control	77	9.5432	0.52418		73	9.2258	0.74646	
T14_M_D_mean	Study	23	6.7028	0.33033	**0.000**	24	6.5146	0.49383	**0.006**
	Control	78	7.1678	0.59534		74	6.8179	0.44571	
T14_B_L_Mean	Study	23	8.8348	0.52633	**0.001**	24	8.6867	0.67602	**0.036**
	Control	78	9.3046	0.57552		74	8.9934	0.59117	
T13_M_D_mean	Study	24	7.6796	0.39377	0.103	23	7.3374	0.35275	0.217
	Control	80	7.8665	0.51295		71	7.4701	0.47038	
T13_B_L_Mean	Study	24	7.6802	0.97182	**0.036**	23	7.4376	0.73175	0.197
	Control	80	8.0754	0.74212		71	7.6647	0.72824	
T12_M_D_mean	Study	24	6.5779	0.44441	**0.009**	24	6.0729	0.77084	**0.01**
	Control	82	6.8951	0.68435		79	6.4771	0.62834	
T12_B_L_Mean	Study	24	5.7477	0.78979	**0.001**	24	5.3448	0.66536	**0.006**
	Control	82	6.3576	0.75832		79	5.8908	0.88094	
T11_M_D_mean	Study	25	8.652	0.46992	**0.037**	25	8.412	0.47384	0.927
	Control	82	8.9068	0.54407		80	8.4241	0.59944	
T11_B_L_Mean	Study	24	6.6648	0.83792	**0.006**	25	6.392	0.50092	**0.001**
	Control	82	7.1647	0.74752		80	6.8332	0.61082	

Table 4.1.2: Independent t-test for MD and BL means of upper left quadrant in male and female group. Significant differences in group means are shown in red.

Statistics of upper left quadrant									
		Male				**Female**			
	Study group	**n**	**Mean**	**Std. Deviation**	**p-value**	**n**	**Mean**	**Std. Deviation**	**P-value**
T21_M_D_mean	Study	25	8.5558	0.63967	**0.006**	25	8.4748	0.47532	0.825
	Control	82	8.9205	0.54945		81	8.4474	0.55913	
T21_B_L_Mean	Study	24	6.5896	0.84664	**0.000**	25	6.4668	0.45074	**0.002**
	Control	82	7.2357	0.74294		81	6.849	0.5564	
T22_M_D_mean	Study	24	6.5319	0.48605	**0.010**	24	6.211	0.64035	0.052
	Control	81	6.8544	0.53611		79	6.5303	0.71227	
T22_B_L_Mean	Study	23	5.7274	0.78878	**0.001**	23	5.4661	0.52106	**0.02**
	Control	81	6.296	0.66822		79	5.8582	0.73948	
T23_M_D_mean	Study	24	7.6925	0.45526	0.277	23	7.1478	0.44712	0.136
	Control	80	7.8228	0.52754		74	7.3241	0.50338	
T23_B_L_Mean	Study	24	7.53	0.79213	**0.013**	23	7.3224	0.70346	0.265
	Control	80	8.0091	0.8172		74	7.5346	0.81873	
T24_M_D_mean	Study	22	6.7859	0.32964	**0.018**	22	6.6209	0.41765	**0.026**
	Control	48	7.0255	0.48012		54	6.8807	0.46479	
T24_B_L_Mean	Study	22	8.7659	0.59635	**0.000**	22	8.6082	0.77327	0.103
	Control	48	9.2317	0.43533		54	8.912	0.70851	
T25_M_D_mean	Study	22	6.593	0.5493	0.201	21	6.3329	0.37046	**0.023**
	Control	49	6.7887	0.60794		51	6.628	0.53103	
T25_B_L_Mean	Study	23	9.0446	0.52585	**0.003**	21	8.7857	0.54044	**0.05**
	Control	79	9.4363	0.55342		71	9.1049	0.67465	
T26_M_D_mean	Study	23	10.0874	0.60904	**0.015**	25	9.8412	0.71558	0.614
	Control	73	10.5095	0.73779		71	9.9174	0.62351	
T26_B_L_Mean	Study	23	11.6252	0.54823	0.665	25	11.1212	0.59445	0.88
	Control	73	11.5523	0.74432		71	11.0965	0.73433	
T27_M_D_mean	Study	15	9.8523	0.47671	**0.052**	18	9.29	0.53571	**0.016**
	Control	61	10.228	0.69444		55	9.7	0.63481	
T27_B_L_Mean	Study	15	11.1667	0.44222	**0.037**	18	10.4508	0.8048	0.09
	Control	61	11.5704	0.69868		55	10.8015	0.73354	

Table 4.1.3: Independent t-test for MD and BL means of lower left quadrant in male and female group. Significant differences in group means are shown in red.

Statistics of lower left quadrant									
		Male				**Female**			
	Study group	**n**	**Mean**	**Std. Deviation**	**P-value**	**n**	**Mean**	**Std. Deviation**	**P-value**
T31_M_D_mean	Study	25	5.3516	0.28152	**0.020**	25	5.2974	0.36704	0.755
	Control	82	5.5315	0.45314		81	5.2716	0.35894	
T31_B_L_Mean	Study	25	5.9678	0.55289	0.064	25	5.6084	0.545	0.136
	Control	82	6.2329	0.63927		81	5.7816	0.49113	
T32_M_D_mean	Study	25	6.0796	0.32654	0.285	25	5.8078	0.33896	0.633
	Control	82	6.1722	0.50349		80	5.7588	0.47475	
T32_B_L_Mean	Study	25	6.1288	0.57423	0.057	25	6.0458	0.56679	0.537
	Control	82	6.3471	0.47024		80	5.9658	0.56335	
T33_M_D_mean	Study	25	6.8354	0.40397	0.256	24	6.2894	0.35476	0.076
	Control	81	6.9678	0.53287		78	6.4544	0.40568	
T33_B_L_Mean	Study	25	7.2758	0.76923	0.193	24	6.8871	0.62044	0.595
	Control	81	7.507	0.77145		78	6.8032	0.6895	
T34_M_D_mean	Study	23	6.8741	0.36669	**0.000**	24	6.61	0.44309	**0.001**
	Control	79	7.3112	0.48374		77	6.9806	0.47078	
T34_B_L_Mean	Study	23	7.4957	0.31909	**0.000**	24	7.309	0.7198	**0.011**
	Control	79	7.9085	0.51194		77	7.6896	0.59692	
T35_M_D_mean	Study	19	7.0682	0.38189	0.061	12	7.0529	0.47034	0.563
	Control	78	7.3054	0.51002		74	7.1621	0.62129	
T35_B_L_Mean	Study	19	8.1908	0.43338	**0.001**	12	7.8275	0.55882	**0.002**
	Control	78	8.7071	0.60384		74	8.4106	0.57509	
T36_M_D_mean	Study	23	10.7726	0.60335	**0.009**	22	10.1859	1.0661	**0.015**
	Control	75	11.2177	0.73116		71	10.7046	0.78901	
T36_B_L_Mean	Study	23	10.498	0.52598	**0.001**	22	10.2102	0.48981	**0.028**
	Control	75	10.906	0.51712		71	10.5587	0.6757	
T37_M_D_mean	Study	17	10.1685	0.51744	**0.029**	17	9.4224	0.46822	**0.000**
	Control	51	10.5522	0.82023		51	10.2793	0.8143	
T37_B_L_Mean	Study	17	10.1938	0.52397	**0.037**	17	9.6838	0.69027	**0.004**
	Control	51	10.5539	0.62582		51	10.2248	0.64005	

Table 4.1.4: Independent t-test for MD and BL means of lower right quadrant in male and female. Significant differences in group means are shown in red.

Statistics of lower right quadrant									
		Male				**Female**			
	Study group	**n**	**Mean**	**Std. Deviation**	**p-value**	**n**	**Mean**	**Std. Deviation**	**P-value**
T41_M_D_mean	Study	24	5.3971	0.34456	0.234	25	5.3304	0.36945	0.17
	Control	82	5.519	0.46201		81	5.1998	0.42469	
T41_B_L_Mean	Study	24	5.8133	0.72711	**0.002**	25	5.754	0.5422	0.535
	Control	82	6.2516	0.56173		81	5.8254	0.488	
T42_M_D_mean	Study	25	5.9398	0.4114	0.154	25	5.7336	0.33224	0.531
	Control	81	6.0851	0.45157		81	5.6767	0.55062	
T42_B_L_Mean	Study	25	6.225	0.63987	**0.078**	25	6.0484	0.63587	0.833
	Control	81	6.4538	0.53517		81	6.0239	0.46304	
T43_M_D_mean	Study	25	6.6626	0.46956	**0.006**	24	6.1473	0.51915	**0.01**
	Control	82	6.963	0.46308		76	6.4164	0.40941	
T43_B_L_Mean	Study	25	7.1676	0.79385	0.107	24	6.9219	0.7213	0.578
	Control	82	7.4558	0.76993		76	6.9999	0.55212	
T44_M_D_mean	Study	22	6.9184	0.50973	**0.018**	24	6.601	0.45942	**0.001**
	Control	79	7.1966	0.4738		78	6.936	0.43007	
T44_B_L_Mean	Study	22	7.6741	0.43043	**0.000**	24	7.4896	0.74446	0.142
	Control	79	8.2159	0.56877		78	7.7356	0.53043	
T45_M_D_mean	Study	18	7.1292	0.51068	0.218	12	6.8179	0.41371	0.339
	Control	79	7.2932	0.50509		73	6.96	0.4829	
T45_B_L_Mean	Study	18	8.275	0.44334	**0.002**	12	8.0392	0.59448	0.083
	Control	79	8.7132	0.72214		73	8.3772	0.62305	
T46_M_D_mean	Study	21	10.7807	0.46135	**0.007**	20	10.4103	0.72501	0.201
	Control	76	11.1806	0.61124		70	10.6421	0.70472	
T46_B_L_Mean	Study	20	10.6828	0.47083	0.324	20	10.2493	0.58299	0.152
	Control	76	10.8382	0.65649		70	10.4916	0.6805	
T47_M_D_mean	Study	17	10.2744	0.42285	**0.051**	15	9.803	0.67731	**0.03**
	Control	54	10.5836	0.85671		51	10.2508	0.69119	
T47_B_L_Mean	Study	17	10.0929	0.71629	**0.001**	15	9.7283	0.66801	**0.001**
	Control	54	10.6475	0.53612		51	10.3593	0.60977	

4.3.3 Comparison of tooth crown measurement variables between the three groups using ANOVA

Male Group

In all the teeth, the mean tooth measurement variable of the study groups (affected and unaffected) in males was smaller than the control group (Tables 4.2.1 to 4.2.4); and in females, the mean tooth measurement variable was also smaller than study groups in all teeth apart from the lower anterior (Tables 4.3.1 to 4.3.4). Analysis with ANOVA showed that in the male group, 21 teeth out of 28 teeth in the study group were smaller ($p \leq 0.05$) in the MD and/or BL widths than the control, but there was no significant difference between the affected and unaffected groups in all teeth. In the female study group, only 15 teeth were significantly smaller ($p \leq 0.05$) in the MD and/or BL widths than the control, but there was no significant difference between the affected and unaffected groups in all teeth. The means of MD dimensions of both the affected and unaffected male groups were significantly smaller than the control group for tooth No. 16, 15, 14 and 34, while for the BL dimensions, the means of both the affected and unaffected male groups were significantly smaller than the control group for tooth No. 15, 12, 21, 24, 34 and 35. The means of the MD dimensions of the affected male group were significantly smaller than the control group for tooth No. 16, 15, 14, 22, 34, 43, and 46; and the means of the BL measurements of the affected male group were significantly smaller than the control group for tooth No. 15, 14, 13, 12, 11, 21, 24, 34 and 35. The means of the MD dimensions of the unaffected male group were significantly smaller than the control group for tooth No. 16, 15, 14, 21, and 34; and the means of the BL dimensions of the unaffected male group were significantly smaller than the control group for tooth No. 15, 12, 21, 22, 24, 25, 34, 35, 36, 41, 44, and 47.

Female Group

In females, the means of the MD dimensions of both the affected and unaffected groups were significantly smaller than the control group for only tooth 37 while the means of the BL dimensions of both the affected and unaffected groups were significantly smaller than the control group for only tooth 11. The means of the MD dimensions of the affected female group were significantly smaller than the control group for tooth No. 15, 14, 12, 25, 27, 34, 36, 37 and 44, whereas the means of the BL dimensions of the affected female group were significantly smaller than the control group for tooth No. 15, 11, 21, 34, 44, 45, and 47.

The means of the MD dimensions of the unaffected female group were significantly smaller than the control group for only tooth 37, and the means of the BL dimensions were significantly smaller than the control group for only tooth No. 37.

Additional Statistical tests

The means of the MD dimensions of tooth 17, 26 and 36 and BL dimensions of tooth No. 26 were significant different among the 3 groups with ANOVA analysis, but the Tukey test did not identify where the significance lay among the three male groups. Thus the Duncan test was used in these teeth and a significant difference was observed between the affected and control groups, but the Duncan test could not identify a significant difference between the groups in the BL dimensions of tooth 23 in the male group and tooth 12 in the female group, nor in the MD dimensions of tooth 43 in the female group. This was because the means of the groups were inhomogeneous in these measurements. The significantly smaller teeth in the male and female groups are illustrated in Figure 4.6. Group means, p values for ANOVA and results for the Post Hoc Tukey HSD tests are shown in Tables 4.2.1 to 4.2.4 and Tables 4.3.1 to 4.3.4.

4.3.4

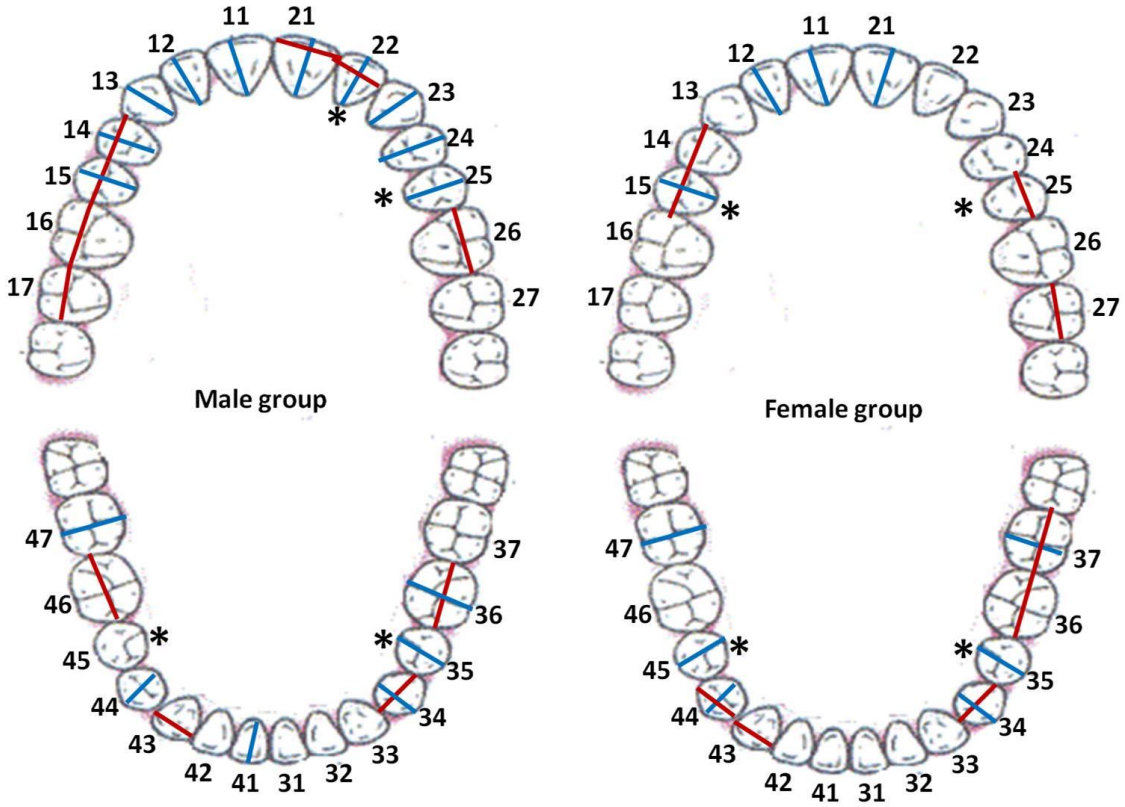


Figure 4.6: The lines indicate that the dimensions of the tooth types that are significantly different ($p \leqslant 0.05$) among the three groups (affected, unaffected and control) in male and female based on ANOVA analysis (red line MD dimension, blue line indicates BL dimension, * indicates congenital missing teeth in the affected group)

Table 4.2.1: Analysis of variance between groups for upper right quadrant for males. Source of variation between groups is shown by Tukey HSD post hoc test. Significant values are shown in red colour. NS - not sig. at level of significance 0.05.

Statistics of upper right quadrant								
Male						Post Hoc Tukey HSD (p value)		
		n	Mean	Std.	ANOVA p-value	Study affected	Study unaffected	Control
T17_M_D_mean	Study affected	5	9.7150	.79892	.01400	1	0.798	0.06*
	Study unaffected	10	9.9560	.52877			1	0.091
	Control	65	10.4518	.69825				1
T17_B_L_mean	Study affected	5	11.2860	.56512	.82700	1	NS	NS
	Study unaffected	10	11.2425	.56346			1	NS
	Control	65	11.3853	.76682				1
T16_M_D_mean	Study affected	8	9.8625	.48137	.00000	1	0.392	0.002
	Study unaffected	17	10.2018	.50483			1	0.002
	Control	76	10.6376	.63152				1
T16_B_L_Mean	Study affected	8	11.1356	.58023	.05700	1	NS	NS
	Study unaffected	17	11.4418	.65655			1	NS
	Control	76	11.6661	.64965				1
T15_M_D_mean	Study affected	7	6.2100	.22264	.00000	1	0.48	0.001
	Study unaffected	17	6.4591	.39662			1	0.003
	Control	77	6.8923	.50749				1
T15_B_L_Mean	Study affected	7	8.6950	.42813	.00000	1	0.155	0
	Study unaffected	17	9.1091	.35416			1	0.004
	Control	77	9.5432	.52418				1
T14_M_D_mean	Study affected	7	6.6007	.39012	.00200	1	0.826	0.028
	Study unaffected	16	6.7475	.30359			1	0.017
	Control	78	7.1678	.59534				1
T14_B_L_Mean	Study affected	7	8.5729	.53702	.00100	1	0.305	0.004
	Study unaffected	16	8.9494	.49489			1	0.06
	Control	78	9.3046	.57552				1
T13_M_D_mean	Study affected	7	7.6779	.51961	.26700	1	NS	NS
	Study unaffected	17	7.6803	.34877			1	NS
	Control	80	7.8665	.51295				1
T13_B_L_Mean	Study affected	7	7.2214	1.11087	.02200	1	0.167	0.02
	Study unaffected	17	7.8691	.87456			1	0.593
	Control	80	8.0754	.74212				1
T12_M_D_mean	Study affected	7	6.6686	.52502	.09900	1	NS	NS
	Study unaffected	17	6.5406	.41888			1	NS
	Control	82	6.8951	.68435				1
T12_B_L_Mean	Study affected	7	5.5986	.74192	.00300	1	0.815	0.036
	Study unaffected	17	5.8091	.82251			1	0.023
	Control	82	6.3576	.75832				1
T11_M_D_mean	Study affected	8	8.6675	.26103	.11400	1	NS	NS
	Study unaffected	17	8.6447	.54886			1	NS
	Control	82	8.9068	.54407				1
T11_B_L_Mean	Study affected	7	6.4271	1.09932	.01500	1	0.596	0.043
	Study unaffected	17	6.7626	.72195			1	0.127
	Control	82	7.1647	.74752				1

***** Tukey test did not identify where the significance lay among the three groups.

Table 4.2.2: Analysis of variance between groups for upper left quadrant for males. Source of variation between groups is shown by Tukey HSD post hoc test. Significant values are shown in red colour. NS - not sig. at level of significance 0.05

Statistics of upper left quadrant								
Male						Post Hoc Tukey HSD (P value)		
		n	Mean	Std.	ANOVA p-value	Study affected	Study unaffected	Control
T21_M_D_mean	Study affected	8	8.7175	.44410	.01500	1	0.597	0.604
	Study unaffected	17	8.4797	.71292			1	0.013
	Control	82	8.9205	.54945				1
T21_B_L_Mean	Study affected	7	6.5036	1.17536	.00200	1	0.934	0.046
	Study unaffected	17	6.6250	.71259			1	0.01
	Control	82	7.2357	.74294				1
T22_M_D_mean	Study affected	7	6.3629	.50139	.02100	1	0.571	0.05
	Study unaffected	17	6.6015	.47716			1	0.173
	Control	81	6.8544	.53611				1
T22_B_L_Mean	Study affected	6	5.6133	.62789	.00300	1	0.888	0.059
	Study unaffected	17	5.7676	.85187			1	0.015
	Control	81	6.2960	.66822				1
T23_M_D_mean	Study affected	7	7.5221	.59069	.32200	1	NS	NS
	Study unaffected	17	7.7626	.38622			1	NS
	Control	80	7.8228	.52754				1
T23_B_L_Mean	Study affected	7	7.4071	.70509	.04100	1	0.884	0.151*
	Study unaffected	17	7.5806	.84038			1	0.125
	Control	80	8.0091	.81720				1
T24_M_D_mean	Study affected	6	6.7592	.32783	.11600	1	NS	NS
	Study unaffected	16	6.7959	.34045			1	NS
	Control	48	7.0255	.48012				1
T24_B_L_Mean	Study affected	6	8.6492	.27332	.00200	1	0.776	0.022
	Study unaffected	16	8.8097	.68226			1	0.011
	Control	48	9.2317	.43533				1
T25_M_D_mean	Study affected	5	6.6680	.76607	.42200	1	NS	NS
	Study unaffected	17	6.5709	.49701			1	NS
	Control	49	6.7887	.60794				1
T25_B_L_Mean	Study affected	6	9.0983	.63913	.01300	1	0.958	0.319
	Study unaffected	17	9.0256	.50110			1	0.017
	Control	79	9.4363	.55342				1
T26_M_D_mean	Study affected	7	9.9393	.51274	.04200	1	0.787*	0.112*
	Study unaffected	16	10.1522	.65127			1	0.169
	Control	73	10.5095	.73779				1
T26_B_L_Mean	Study affected	7	11.5986	.33729	.90500	1	NS	NS
	Study unaffected	16	11.6369	.62836			1	NS
	Control	73	11.5523	.74432				1
T27_M_D_mean	Study affected	4	9.7825	.66124	.14800	1	NS	NS
	Study unaffected	11	9.8777	.42932			1	NS
	Control	61	10.2280	.69444				1
T27_B_L_Mean	Study affected	4	11.1600	.65898	.11400	1	NS	NS
	Study unaffected	11	11.1691	.37879			1	NS
	Control	61	11.5704	.69868				1

***** Tukey test did not identify where the significance lay among the three groups.

Table 4.2.3: Analysis of variance between groups for lower left quadrant for males. Source of variation between groups is shown by Tukey HSD post hoc test. Significant values are shown in red colour. NS - not sig. at level of significance 0.05.

Statistics of lower left quadrant								
Male						Post Hoc Tukey HSD (P value)		
		n	Mean	Std.	ANOVA p-value	Study affected	Study unaffected	Control
T31_M_D_mean	Study affected	8	5.3250	.33966	.17600	1	NS	NS
	Study unaffected	17	5.3641	.26054			1	NS
	Control	82	5.5315	.45314				1
T31_B_L_Mean	Study affected	8	6.0900	.44707	.14500	1	NS	NS
	Study unaffected	17	5.9103	.60009			1	NS
	Control	82	6.2329	.63927				1
T32_M_D_mean	Study affected	8	6.1269	.27059	.65200	1	NS	NS
	Study unaffected	17	6.0574	.35534			1	NS
	Control	82	6.1722	.50349				1
T32_B_L_Mean	Study affected	8	6.0581	.67734	.14600	1	NS	NS
	Study unaffected	17	6.1621	.53872			1	NS
	Control	82	6.3471	.47024				1
T33_M_D_mean	Study affected	8	6.6569	.53678	.25300	1	NS	NS
	Study unaffected	17	6.9194	.30869			1	NS
	Control	81	6.9678	.53287				1
T33_B_L_Mean	Study affected	8	7.1969	.84949	.40400	1	NS	NS
	Study unaffected	17	7.3129	.75317			1	NS
	Control	81	7.5070	.77145				1
T34_M_D_mean	Study affected	8	6.8619	.28633	.00100	1	0.995	0.028
	Study unaffected	15	6.8807	.41253			1	0.004
	Control	79	7.3112	.48374				1
T34_B_L_Mean	Study affected	8	7.4588	.31366	.00200	1	0.961	0.034
	Study unaffected	15	7.5153	.33108			1	0.012
	Control	79	7.9085	.51194				1
T35_M_D_mean	Study affected	3	7.4233	.41816	.06800	1	NS	NS
	Study unaffected	16	7.0016	.34890			1	NS
	Control	78	7.3054	.51002				1
T35_B_L_Mean	Study affected	3	7.8333	.24420	.00200	1	0.471	0.03
	Study unaffected	16	8.2578	.43253			1	0.015
	Control	78	8.7071	.60384				1
T36_M_D_mean	Study affected	8	10.6244	.61941	.02600	1	0.743*	0.066*
	Study unaffected	15	10.8517	.60080			1	0.164*
	Control	75	11.2177	.73116				1
T36_B_L_Mean	Study affected	8	10.3569	.78218	.00400	1	0.609	0.066
	Study unaffected	15	10.5733	.33373			1	0.015
	Control	75	10.9060	.51712				1
T37_M_D_mean	Study affected	5	10.2550	.59932	.19900	1	NS	NS
	Study unaffected	12	10.1325	.50401			1	NS
	Control	51	10.5522	.82023				1
T37_B_L_Mean	Study affected	5	10.1140	.64878	.10800	1	NS	NS
	Study unaffected	12	10.2271	.49211			1	NS
	Control	51	10.5539	.62582				1

***** Tukey test did not identify where the significance lay among the three groups.

Table 4.2.4: Analysis of variance between groups for lower right quadrant for males. Source of variation between groups is shown by Tukey HSD post hoc test. Significant values are shown in red colour. NS - not sig. at level of significance 0.05.

Statistics of lower right quadrant								
Male						Post Hoc Tukey HSD (P value)		
		n	Mean	Std.	ANOVA P-value	Study affected	Study unaffected	Control
T41_M_D_mean	Study affected	8	5.3588	.44409	.47200	1	NS	NS
	Study unaffected	16	5.4163	.29805			1	NS
	Control	82	5.5190	.46201				1
T41_B_L_Mean	Study affected	8	5.9294	.88804	.00800	1	0.784	0.324
	Study unaffected	16	5.7553	.65715			1	0.009
	Control	82	6.2516	.56173				1
T42_M_D_mean	Study affected	8	5.9319	.46842	.36400	1	0.998	0.623
	Study unaffected	17	5.9435	.39729			1	0.460
	Control	81	6.0851	.45157				1
T42_B_L_Mean	Study affected	8	6.2663	.77431	.20600	1	NS	NS
	Study unaffected	17	6.2056	.59211			1	NS
	Control	81	6.4538	.53517				1
T43_M_D_mean	Study affected	8	6.4288	.59560	.00500	1	0.194	0.006
	Study unaffected	17	6.7726	.36787			1	0.271
	Control	82	6.9630	.46308				1
T43_B_L_Mean	Study affected	8	6.8356	.99694	.09300	1	NS	NS
	Study unaffected	17	7.3238	.65531			1	NS
	Control	82	7.4558	.76993				1
T44_M_D_mean	Study affected	7	6.8493	.28048	.05700	1	NS	NS
	Study unaffected	15	6.9507	.59373			1	NS
	Control	79	7.1966	.47380				1
T44_B_L_Mean	Study affected	7	7.7236	.42226	.00000	1	0.954	0.062
	Study unaffected	15	7.6510	.44687			1	0.001
	Control	79	8.2159	.56877				1
T45_M_D_mean	Study affected	1	7.3900		.40800	1	NS	NS
	Study unaffected	17	7.1138	.52210			1	NS
	Control	79	7.2932	.50509				1
T45_B_L_Mean	Study affected	1	7.8350		.04400	1		
	Study unaffected	17	8.3009	.44274			1	
	Control	79	8.7132	.72214				1
T46_M_D_mean	Study affected	7	10.6200	.59903	.01700	1	0.646	0.044
	Study unaffected	14	10.8611	.37547			1	0.149
	Control	76	11.1806	.61124				1
T46_B_L_Mean	Study affected	6	10.6233	.67164	.59300	1	NS	NS
	Study unaffected	14	10.7082	.38492			1	NS
	Control	76	10.8382	.65649				1
T47_M_D_mean	Study affected	6	10.1600	.24370	.33600	1	NS	NS
	Study unaffected	11	10.3368	.49422			1	NS
	Control	54	10.5836	.85671				1
T47_B_L_Mean	Study affected	6	10.3483	.33089	.00200	1	0.377	0.458
	Study unaffected	11	9.9536	.84005			1	0.002
	Control	54	10.6475	.53612				1

Table 4.3.1: Analysis of variance between female groups for upper right quadrant for females. Source of variation between groups is shown by Tukey HSD post hoc test. Significant values are shown in red colour. NS - not sig. at level of significance 0.05.

Statistics of upper right quadrant								
Female						Post Hoc Tukey HSD (P value)		
		n	Mean	Std.	ANOVA P-value	Study affected	Study unaffected	Control
T17_M_D_mean	Study affected	8	9.0888	.49761	.06500	1	NS	NS
	Study unaffected	9	9.4850	.45793			1	NS
	Control	59	9.5936	.58535				1
T17_B_L_mean	Study affected	8	10.5731	.67672	.34300	1	NS	NS
	Study unaffected	9	10.4372	.76079			1	NS
	Control	59	10.7900	.74014				1
T16_M_D_mean	Study affected	13	9.8750	.66822	.66600	1	NS	NS
	Study unaffected	11	9.8877	.48135			1	NS
	Control	72	10.0108	.61726				1
T16_B_L_Mean	Study affected	13	11.1746	.70687	.37400	1	NS	NS
	Study unaffected	11	10.9877	.47932			1	NS
	Control	72	11.2563	.59315				1
T15_M_D_mean	Study affected	6	5.9192	.50877	.00800	1	0.033	0.005
	Study unaffected	11	6.5300	.30743			1	0.974
	Control	73	6.5632	.48816				1
T15_B_L_Mean	Study affected	6	8.4050	.69309	.03100	1	0.087	0.024
	Study unaffected	11	9.1886	.51478			1	0.986
	Control	73	9.2258	.74646				1
T14_M_D_mean	Study affected	12	6.3200	.48012	.00300	1	0.091	0.002
	Study unaffected	12	6.7092	.44359			1	0.718
	Control	74	6.8179	.44571				1
T14_B_L_Mean	Study affected	12	8.5604	.79160	.06700	1	NS	NS
	Study unaffected	12	8.8129	.54235			1	NS
	Control	74	8.9934	.59117				1
T13_M_D_mean	Study affected	11	7.3600	.42198	.45600	1	NS	NS
	Study unaffected	12	7.3167	.29326			1	NS
	Control	71	7.4701	.47038				1
T13_B_L_Mean	Study affected	11	7.6923	.48617	.12000	1	NS	NS
	Study unaffected	12	7.2042	.85543			1	NS
	Control	71	7.6647	.72824				1
T12_M_D_mean	Study affected	12	5.9533	.97303	.02600	1	0.653	0.033
	Study unaffected	12	6.1925	.51422			1	0.354
	Control	79	6.4771	.62834				1
T12_B_L_Mean	Study affected	12	5.3325	.76612	.02400	1	.997*	.086*
	Study unaffected	12	5.3571	.58171			1	.106*
	Control	79	5.8908	.88094				1
T11_M_D_mean	Study affected	13	8.4285	.57156	.98500	1	NS	NS
	Study unaffected	12	8.3942	.36445			1	NS
	Control	80	8.4241	.59944				1
T11_B_L_Mean	Study affected	13	6.3700	.33537	.00600	1	0.979	0.027
	Study unaffected	12	6.4158	.65083			1	0.062
	Control	80	6.8332	.61082				1

***** Tukey test did not identify where the significance lay among the three groups.

Table 4.3.2: Analysis of variance between groups for upper left quadrant for females. Source of variation between groups is shown by Tukey HSD post hoc test. Significant values are shown in red colour. NS - not sig. at level of significance 0.05.

Statistics of upper left quadrant								
Female						Post Hoc Tukey HSD (P value)		
		n	Mean	Std.	ANOVA P-value	Study affected	Study unaffected	Control
T21_M_D_mean	Study affected	13	8.4296	.59515	.88900	1	NS	NS
	Study unaffected	12	8.5238	.31858			1	NS
	Control	81	8.4474	.55913				1
T21_B_L_Mean	Study affected	13	6.4419	.39516	.00900	1	0.968	0.033
	Study unaffected	12	6.4938	.52096			1	0.087
	Control	81	6.8490	.55640				1
T22_M_D_mean	Study affected	12	6.1113	.73506	.12000	1	NS	NS
	Study unaffected	12	6.3108	.54345			1	NS
	Control	79	6.5303	.71227				1
T22_B_L_Mean	Study affected	11	5.5273	.45670	.06100	1	NS	NS
	Study unaffected	12	5.4100	.58841			1	NS
	Control	79	5.8582	.73948				1
T23_M_D_mean	Study affected	11	7.2436	.49917	.22200	1	NS	NS
	Study unaffected	12	7.0600	.39462			1	NS
	Control	74	7.3241	.50338				1
T23_B_L_Mean	Study affected	11	7.3786	.60431	.51200	1	NS	NS
	Study unaffected	12	7.2708	.80726			1	NS
	Control	74	7.5346	.81873				1
T24_M_D_mean	Study affected	10	6.5110	.43864	.05000	1	NS	NS
	Study unaffected	12	6.7125	.39427			1	NS
	Control	54	6.8807	.46479				1
T24_B_L_Mean	Study affected	10	8.4375	.98611	.16200	1	NS	NS
	Study unaffected	12	8.7504	.54532			1	NS
	Control	54	8.9120	.70851				1
T25_M_D_mean	Study affected	9	6.0794	.40010	.00900	1	0.097	0.006
	Study unaffected	12	6.5229	.20289			1	0.774
	Control	51	6.6280	.53103				1
T25_B_L_Mean	Study affected	9	8.5739	.44212	.06400	1	NS	NS
	Study unaffected	12	8.9446	.56977			1	NS
	Control	71	9.1049	.67465				1
T26_M_D_mean	Study affected	13	9.9827	.76002	.46400	1	NS	NS
	Study unaffected	12	9.6879	.66165			1	NS
	Control	71	9.9174	.62351				1
T26_B_L_Mean	Study affected	13	11.2396	.69140	.67400	1	NS	NS
	Study unaffected	12	10.9929	.46367			1	NS
	Control	71	11.0965	.73433				1
T27_M_D_mean	Study affected	10	9.1260	.42814	.02500	1	0.414	0.021
	Study unaffected	8	9.4950	.61224			1	0.649
	Control	55	9.7000	.63481				1
T27_B_L_Mean	Study affected	10	10.2880	.96156	.14200	1	NS	NS
	Study unaffected	8	10.6544	.54681			1	NS
	Control	55	10.8015	.73354				1

Table 4.3.3: Analysis of variance between groups for lower left quadrant for females. Source of variation between groups is shown by Tukey HSD post hoc test. Significant values are shown in red colour. NS - not sig. at level of significance 0.05.

Statistics of lower left quadrant								
Female						Post Hoc Tukey HSD (P value)		
		n	Mean	Std.	ANOVA P-value	Study affected	Study unaffected	Control
T31_M_D_mean	Study affected	13	5.2565	.31348	.80200	1	NS	NS
	Study unaffected	12	5.3417	.42733			1	NS
	Control	81	5.2716	.35894				1
T31_B_L_Mean	Study affected	13	5.6185	.44158	.32900	1	NS	NS
	Study unaffected	12	5.5975	.65962			1	NS
	Control	81	5.7816	.49113				1
T32_M_D_mean	Study affected	13	5.7715	.36655	.81700	1	NS	NS
	Study unaffected	12	5.8471	.31759			1	NS
	Control	80	5.7588	.47475				1
T32_B_L_Mean	Study affected	13	6.1212	.53676	.65100	1	NS	NS
	Study unaffected	12	5.9642	.61043			1	NS
	Control	80	5.9658	.56335				1
T33_M_D_mean	Study affected	12	6.3050	.41498	.20500	1	NS	NS
	Study unaffected	12	6.2738	.30069			1	NS
	Control	78	6.4544	.40568				1
T33_B_L_Mean	Study affected	12	6.9038	.57480	.86300	1	NS	NS
	Study unaffected	12	6.8704	.68840			1	NS
	Control	78	6.8032	.68950				1
T34_M_D_mean	Study affected	12	6.4842	.55335	.00200	1	0.381	0.002
	Study unaffected	12	6.7358	.26414			1	0.209
	Control	77	6.9806	.47078				1
T34_B_L_Mean	Study affected	12	7.0979	.89876	.01000	1	0.225	0.008
	Study unaffected	12	7.5200	.42237			1	0.655
	Control	77	7.6896	.59692				1
T35_M_D_mean	Study affected	1	7.1650		.83100	1	NS	NS
	Study unaffected	11	7.0427	.49190			1	NS
	Control	74	7.1621	.62129				1
T35_B_L_Mean	Study affected	1	7.5150		.00600	1		
	Study unaffected	11	7.8559	.57694			1	
	Control	74	8.4106	.57509				1
T36_M_D_mean	Study affected	11	9.9418	1.39179	.02200	1	0.379	0.026
	Study unaffected	11	10.4300	.56449			1	0.586
	Control	71	10.7046	.78901				1
T36_B_L_Mean	Study affected	11	10.3455	.58653	.05500	1	NS	NS
	Study unaffected	11	10.0750	.34579			1	NS
	Control	71	10.5587	.67570				1
T37_M_D_mean	Study affected	9	9.5206	.49997	.00100	1	0.835	0.018
	Study unaffected	8	9.3119	.43482			1	0.003
	Control	51	10.2793	.81430				1
T37_B_L_Mean	Study affected	9	9.7789	.73443	.01400	1	0.802	0.152
	Study unaffected	8	9.5769	.66928			1	0.031
	Control	51	10.2248	.64005				1

Table 4.3.4: Analysis of variance between groups for lower right quadrant for females. Source of variation between groups is shown by Tukey HSD post hoc test. Significant values are shown in red colour. NS - not sig. at level of significance 0.05.

Statistics of lower right quadrant								
Female						Post Hoc Tukey HSD (P value)		
		n	Mean	Std.	ANOVA P-value	Study affected	Study unaffected	Control
T41_M_D_mean	Study affected	13	5.3158	.41016	.38400	1	NS	NS
	Study unaffected	12	5.3463	.33727			1	NS
	Control	81	5.1998	.42469				1
T41_B_L_Mean	Study affected	13	5.8238	.60164	.63600	1	NS	NS
	Study unaffected	12	5.6783	.48429			1	NS
	Control	81	5.8254	.48800				1
T42_M_D_mean	Study affected	13	5.7285	.34837	.88700	1	NS	NS
	Study unaffected	12	5.7392	.32922			1	NS
	Control	81	5.6767	.55062				1
T42_B_L_Mean	Study affected	13	6.0715	.70923	.95100	1	NS	NS
	Study unaffected	12	6.0233	.57631			1	NS
	Control	81	6.0239	.46304				1
T43_M_D_mean	Study affected	12	6.1604	.59859	.03700	1	0.988*	0.152*
	Study unaffected	12	6.1342	.45260			1	0.102*
	Control	76	6.4164	.40941				1
T43_B_L_Mean	Study affected	12	6.9463	.87224	.84000	1	NS	NS
	Study unaffected	12	6.8975	.57073			1	NS
	Control	76	6.9999	.55212				1
T44_M_D_mean	Study affected	12	6.5558	.53405	.00600	1	0.869	0.017
	Study unaffected	12	6.6463	.38942			1	0.089
	Control	78	6.9360	.43007				1
T44_B_L_Mean	Study affected	12	7.2758	.92989	.04200	1	0.173	0.032
	Study unaffected	12	7.7033	.44097			1	0.982
	Control	78	7.7356	.53043				1
T45_M_D_mean	Study affected	2	6.4500	.30406	.30900	1	NS	NS
	Study unaffected	10	6.8915	.40351			1	NS
	Control	73	6.9600	.48290				1
T45_B_L_Mean	Study affected	2	7.1575	1.00056	.01900	1	0.067	0.017
	Study unaffected	10	8.2155	.33677			1	0.708
	Control	73	8.3772	.62305				1
T46_M_D_mean	Study affected	11	10.3086	.71767	.34500	1	NS	NS
	Study unaffected	9	10.5344	.75699			1	NS
	Control	70	10.6421	.70472				1
T46_B_L_Mean	Study affected	11	10.4382	.68288	.13200	1	NS	NS
	Study unaffected	9	10.0183	.33946			1	NS
	Control	70	10.4916	.68050				1
T47_M_D_mean	Study affected	10	9.8360	.60906	.09400	1	NS	NS
	Study unaffected	5	9.7370	.87337			1	NS
	Control	51	10.2508	.69119				1
T47_B_L_Mean	Study affected	10	9.5905	.57133	.00200	1	0.448	0.002
	Study unaffected	5	10.0040	.82758			1	0.445
	Control	51	10.3593	.60977				1

***** Tukey test did not identify where the significance lay among the three groups.

4.4 Discussions

Many studies have revealed smaller tooth dimensions of the present teeth in individuals with tooth agenesis. Some of them measured only the mesio-distal dimension (Garn and Lewis, 1970; Baum and Cohen, 1971; Rune and Sarnas, 1974). More recently, Brook *et al.*, reported that patients with hypodontia tended to have smaller remaining teeth than a control group and this tendency increased with increase in the number of missing teeth (Brook et al., 2009b). However there were a few studies that investigated families with tooth agenesis and revealed a reduction in tooth size in male and female relatives of patients with tooth agenesis. (Schalkvanderweide *et al.,* 1994; McKeown *et al.,* 2002; Brook *et al.,* 2009a; Brook *et al.,* 2009b). Their results demonstrated a pattern of a similar and greater reduction in bucco-lingual dimensions compared with mesio-distal dimensions.

As mentioned in the literature review, both genetic and environmental factors can cause reduction in tooth size (Lussi and Jaeggi, 2008, Brook *et al*., 2009a). In tooth size measurement study, age and environmental factors such as diet causing erosion and/or abrasion should be considered as confounding factors. In this study, the control group was selected from the same ethnic population with similar age range in order to minimise these factors.

In this study MD and BL were measured manually and reductions in both dimensions were noticed almost with the same number of teeth, but there was variability with regard to the tooth type, as has been reported in other studies (McKeown *et al.,* 2002; Brook *et al.,* 2009a). Although the manual method can offer accurate and reproducible measures, the number of measurements that can be obtained are limited. Also, it is a time consuming method and subjective error is possible with this technique. However there have been other studies that have investigated tooth dimensions using image analysis systems and reported almost the

same results of reductions in tooth size from different views such as the buccal and occlusal views (Brook *et al.,* 1986). The main advantages of this method over the manual method are the ability to make multiple measurements from a single image and that the automation of the measurement procedures may decrease subjectivity in identifying landmarks. However, identification of appropriate landmarks could also be subjective. New three-dimensional imaging has the advantage of providing a set of anatomical data that can be used to study overall effect of mutated gene on teeth and palatal vault. However, recording undercut regions is an issue with this technique. Nevertheless, all the studies concluded that there is a reduction in tooth dimensions for people with hypodontia. A study on the tooth dimensions, measured by image analysis, in a family with severe tooth agenesis having a PAX9 gene mutation revealed a significant reduction in tooth crown size in all teeth apart from the upper left second premolars on the MD dimension (Brook *et al.,* 2009b).

In a study by Mckeown *et al.,* (2002) of 12 families with a severe type of tooth agenesis (missing six teeth or more), the authors studied the tooth dimensions of the family index members and their relatives and found that the canines and first premolars in both arches, upper central incisors, upper first molars, lower lateral incisors and second premolars were significantly smaller in all dimensions from the occlusal view. For the relatives group, the first premolars were significantly smaller for all dimensions from the occlusal view and upper second premolars were significantly smaller in MD from the buccal view. Although Mckeown and his coworkers did not specify that the 12 families had the same phenotype with regard to tooth type, nor the causative mutations in these families, it has been proposed that every gene involved in tooth development may have its specific function and different role for all the teeth to produce different phenotypes of tooth agenesis and reduction in tooth size.

Studying these tooth dimensions has helped researchers to clarify the range of tooth crown reduction in individuals with tooth agenesis.

Family members with tooth agenesis in *Family I* had smaller tooth sizes than the control group. This finding is consistent with previous studies which reported a link between tooth agenesis and reduction in the tooth crown size. These studies indicated that this link may be due to genetic control. To clarify this, *Family I* members were considered as one unit and compared to the control group by independent t- test. This type of analysis was carried out for two reasons, first to investigate the assumption that this family might have smaller teeth than the control as part of their genetic features such as body build, skull size and shape of the face. The second was to improve the sample size based on arguments concerning the independency of the affected members and their relatives. Thus these two groups were considered as one group. The analysis of the tooth crown dimensional means (MD and BL) with independent t tests in the present study showed that in the male group, the MD and BL means of almost all upper teeth and most of the lower teeth were significantly smaller than the control (Tables 4.1.1 to 4.1.4). The female group showed less difference in the dimensions of their teeth relative to their controls than the male group when the significant differences were mainly found in the posterior teeth. This observation may be related to the smaller size of normal female teeth relative to normal male teeth, thus, the power to attain a significant difference is reduced. However, this observation has also been noticed by another group (Brook et al., 2009b). The independent t test analysis for *Family I* confirmed that their teeth were significantly smaller than the control group from the same region. This may be due to influential impact of the tooth agenesis mutated gene(s) on the region of tooth agenesis or it may be generally genetic make-up of this family having smaller teeth.

When the individuals in this family were divided into two groups, affected and unaffected, in the second part of this study, the ANOVA test applied to the three groups revealed that there were no significant divergences in the MD and BL variables between the groups for the affected and their unaffected relatives. These results support the assumption in the first part that the affected and their unaffected relatives could be considered as one group, sharing common genetic factor or factors. Although the present study showed that the MD and BL dimensions of some teeth in the affected group were reduced in individuals with tooth agenesis compared to their unaffected relatives and the control group, the pattern of the reduction in tooth crown size in the male and female groups was asymmetric (Figure 4.6). There was no obvious pattern in the variability of tooth dimensions which can be attributed to tooth type or time of development. This suggests that there are influences from the other genes controlling tooth development reciprocally (Thesleff, 2006), and/or having gene compensating role (Tucker and Sharpe, 2004).

The ANOVA analysis demonstrated that although the tooth crown sizes of the affected and unaffected groups were smaller than that of the control, the difference is inconsistent with regard to the tooth type (Tables 4.2.1 to 4.2.4) and (Tables 4.3.1 to 4.3.4). This inconsistency, as mentioned before, may be due to the relatively small sample size in only one family. A study involving more affected families to increase statistical power is needed to investigate the association of tooth agenesis and small size tooth type. With a larger sample size, the analysis can be stratified into the number of tooth agenesis according to their tooth type, rather than grouping them together as in the present study. However, a modified version of the conventional analysis of variance (ANOVA) called a Procrustes ANOVA has been introduced (Klingenberg *et al.*, 2002) to assess the relative magnitudes of measurements error from repeat measurements in geometric morphometrics studies of two and three land mark

data. The Procrustes ANOVA could be used in the future work to assess the tooth size measurements in the Saudi family.

4.5 Summary

The measurement of tooth crown dimensions of all teeth in the affected group and their relatives revealed a trend of reduction in the crown sizes relative to the control group. There was no significant difference between the tooth crown sizes of the affected and their relatives. This supports the suggestion that there may be an interaction between the genes responsible for tooth agenesis and the gene for tooth size; or that the gene for tooth agenesis also controls tooth size with varying degrees of penetrance. Apart from sizes, there may be association with other features such as shapes of the teeth and arch dimensions indicating the interaction of the agenesis gene with other genes such as the morphogenic genes. The limitation of the present study is that only one family, though a large cohort, was involved and the phenotype for selection is missing second premolars and upper lateral incisors because they are the commonest missing tooth types. The challenge is to find more families with large pedigrees like this family with different missing tooth types to confirm the hypothesis.

CHAPTER 5: OVERALL DISCUSSION, STUDIES' LIMITATIONS AND FUTURE WORK

5.1 Final discussion

This is the first study of tooth agenesis to investigate the phenotype and genotype of non syndromic tooth agenesis in Saudi population. In this population consanguineous marriage is common suggesting a relatively pure genetic pool with the potential of locating autosomal recessive gene. Based on this hypothesis the study was designed hunting for novel mutations that can be linked to phenotype of tooth agensis.

As mentioned in the literature review, the proper clinical assessment of the families phenotypes is the gate that leads to the identification of the responsible gene and its function. Even though the researchers cannot link the phenotype to the causative mutation or responsible variants in the present time, this valuable information will help to explore the tooth agenesis variable features and their molecular basis in the future.

In this large study, 16 families were reported, and each family was studied as an independent case with regard to the pattern of inheritance, number of tooth types missing, association of other dental anomalies and number of affected members in the families. These factors were all considered in the selection of the families for molecular investigations.

The data in this study might not be sufficient to represent the prevalence of the general phenotype of familial non syndromic tooth agenesis in Saudi population, but it is enough to highlight the nature of this dental condition in this population by showing different phenotypes with different tooth type affected and different modes of inheritance of tooth agenesis in the families. By comparing these phenotypes to the Pakistani family which was also from a consanguineous population, the decision was made to select the suitable families for the whole exome sequencing.

It is interesting to note that in all 16 families the most affected tooth type were the lower second premolars and upper lateral incisors, which are the predicted missing teeth types in according to the hypothesises of the Butler's theory and anatomical model (Bailit, 1975, Svinhufvud *et al.*, 1988).

The results of this study show that in *Family II* , a heterozygous novel frameshift mutation of the MSX1 gene was linked to the agenesis of only the second premolars and third molars. This finding is also consistent with the Butler's theory. In *Family III a* homozygous novel mutation of the SMOC2 gene was linked to the agenesis of premolars, canines and lateral incisors which seems to fit to the anatomical model. These observations suggest that the involved genes in these phenotypes might work in a frame that similar to what have been proposed in these theories.

One of the issue that has been faced in this study is the pedigree analysis of the families. Some families was showing a confusing pattern of the disease inheritance in their pedigree by showing an autosomal dominant pattern in one branch and autosomal recessive pattern in other branches as observed in *Family I* and *FamilyVII.* This can be explained either by the complexity of the disease or the variability in the penetrance of the disease mutation.

The data on the clinical features of those families gave more power to the study as it was valuable in gaining an understanding about every phenotype and in searching for the variability of gene expression in the families with identified mutations. Nieminen has proposed that available evidence about the extensive genetic heterogeneity of tooth agenesis suggests that common types of tooth agenesis may follow different patterns of inheritance even if most of the reported studies have demonstrated an autosomal mode of inheritance (Nieminen, 2009). In that study, 13 families, with different severity in their phenotypes segregated the disease in an autosomal recessive pattern and two were sporadic and only one presented with an autosomal dominant pattern. This suggests that the concentration of

researchers on the autosomal dominant families with affected members in all generations has a bias effect on general knowledge about tooth agenesis. Thus more case reports about different phenotypes of tooth agenesis should be supported as this is the most effective way available to identify the role of new genes in tooth agenesis in human.

With more cases, this may help in providing a better understanding of the wide spectrum of tooth agenesis, as noticed by several researchers that there is a broad range of variability among affected members in the same families, and among families with different types of tooth agenesis. This variability in phenotypes requires researchers and clinicians to continue investigations and research on more clinical cases of families with tooth agenesis, as every case might reflect another function, or role, of a new or a previously unknown gene. This would enables researchers in the dental and medical fields to explore more on the reciprocal roles of different genes in controlling different parts of the human body. The results of different part of the study presented in this thesis on tooth agenesis in consanguineous families have supported most of the observations and findings about tooth agenesis in the literature and added more information on the molecular basis of tooth agenesis through identification of two novel mutations in two families and exclusion of the role of other genes which were linked to other non-syndromic tooth agenesis conditions. This might suggest involvement of other unreported genes in the agenesis of different types of tooth.

As described in Chapter 3 a molecular analysis was carried out on eight families; the molecular basis of tooth agenesis was identified only in two families by exome sequencing. Contrary to expectations, the homozygosity mapping approach has failed to identify the regions of the tooth agenesis genes in the Saudi families. There are several possible explanations for this failure; it is possible that the responsible gene for the tooth agenesis phenotype was different in each family. It is also can be explained by the fact that the

responsible gene(s) in these families might be in one of the suggested regions but it is still unknown or has not been annotated yet. Another explanation could be that in all the families the phenotype was mild and caused by the aggregation of more than one gene defect (complex model).

Microsatellite marker mapping has been used in the beginning of this study to investigate some interested regions, and a concern has been raised about the suitability of the microsatellite marker set that has been used for genome wide linkage scan to the study population (Saudi population). For this reason this approach has been discontinue. Ideally a new set of microsatellite marker specific to Saudi or Arab population should be used.

All genes that have been linked to tooth agenesis (Table 1.1) and their roles in tooth development confirmed their important roles in the formation and regulation of other tissues and organs in the human body. Up to now, not even a single gene has been reported to be involved only in tooth formation (Thesleff, 2006) and not involved in regulation of other tissues, *e.g.* EDA and AXIN2 genes. Thus tooth agenesis researchers might consider changing the trend of dividing tooth agenesis to non-syndromic and syndromic. When they investigate the molecular basis of this condition, they should accept the concept of the unity of the human body (Lammi *et al.,* 2004a; Zhang *et al.,* 2009) as there are clearly other related disease genes to be discovered. For example, a Pakistani family with an autosomal condition of tooth malformation, enamel hypoplasia and failure of eruption of teeth has been mapped to chromosome 16q12.1 but the causative gene mutation has yet to be identified (Ahmad *et al.,* 1998). A severe tooth agenesis phenotype known as He-Zaho deficit was found in a large Chinese family kinship living in a small village in China, where 52 out of 328 family members of six different generations were segregating the disease in an autosomal dominant pattern with incomplete penetrance and high variability in the type and number of missing teeth that range from few teeth missing to anodontia. The disease was mapped to

chromosome 10q11.2, the locus contain many candidates genes that have been suggested as responsible for generating the He-Zaho deficit such as DKK-1, PRKG1B, ZNF11, ZNF22, and ZNF25 but none of them have yet been confirmed (Wang *et al.,* 2000; Liu *et al.,* 2001).

In *Family I* an interesting observation was made among the affected male and female members as some of them had bilateral tooth agenesis of the lower second premolars and others had a unilateral lower second premolar agenesis either in the right or the left. This finding supports the epigenetic influences hypothesis. Townsend *et al.,* in a twin study concluded that the variable phenotypes of tooth agenesis in identical twins cannot be completely explained just by their genotypes. They suggested that phenotypes might be modified by the effects of slight variations on the self-organising developmental process (Townsend *et al.,* 2005).

In chapter 4 an initial study on tooth crown sizes of *Family I* with tooth agenesis investigated the hypothesis of the polygenic model which was proposed by Brook in 1984 and agreed by other researches who studied the relations between reduction in tooth size and presence of missing teeth by examining the teeth sizes of the affected members and their relatives.(Brook, 1984; SchalkvanderWeide *et al.,* 1994; SchalkvanderWeide and Bosman, 1996; McKeown *et al.,* 2002; Brook *et al.,* 2009a; Brook *et al.,* 2009b). The preliminary data analysis supported the association between the tooth agenesis and the reduction in size of the remaining dentition. However the molecular study on mice and reported cases in humans showed variability in the expressions of the genes controlling tooth development; this fact has a direct influence on the studies that have been done to investigate the tooth sizes of individuals with tooth agenesis from different families as some of those families might segregate in to different mutations of different teeth with different roles in tooth development. In these types of studies, subjects should be selected based on the causative gene rather than the type or number of teeth missing.

5.2 Limitations of the studies

5.2.1 Limitations of the exome sequencing study

In exome sequencing, the coverage of regions of interest is not complete and there are some missed areas; it will never reach 100% coverage. The previous limitation can be combated by the synergy of the homzygosity mapping and exome sequencing. Several disease mutations have been identified by this strategy (Theis *et al.*, 2011).

- Up to now exome sequencing is not able to identify repeat mutations(Singleton, 2011).
- The information available for all truly protein-coding exons in the genome is still incomplete, and current capture probes can only target exons that have been identified so far.
- The exome sequencing templates cannot all be synthesised with the same efficiency, and not all sequences can be aligned to the reference genome.

5.2.2 Limitation of the tooth crown dimensions study

The tooth crown dimensions study was needed in this dissertation to assess the expression of the agenesis gene on the entire dentition.

- The study number of recruited participant was relatively low, affected group sample was (8 males and 13 females) and affected relatives group sample was (17 male and 12 female). However previous published studies included a sample of 10 (Brook *et al*., 2009b) and 12 (McKeown *et al*., 2002) members only.
- In this study the number of crown dimensions measurements per tooth was only two measurements (one measurements and repeated measurements). This has been considered as the lower limit since it is commonly believed that the mean of a series of repeated measurements is the best method to estimate the object's true size.

Nevertheless technical errors of measurement are common in this type of study, *e.g.* measurements on dental cast are systemically smaller mesio-distally than measurements of isolated teeth, extracted teeth or when teeth are in situ, because the impression material does not preserve the infinitesimal space between tightly approximated teeth.

- In this study the number of crown dimensions measurements per tooth was only two measurements (one measurements and repeated measurements). This has been considered as the lower limit since it is commonly believed that the mean of a series of repeated measurements is the best method to estimate the object's true size. Nevertheless technical errors of measurement are common in this type of study, *e.g.* measurements on dental cast are systemically smaller mesio-distally than measurements of isolated teeth, extracted teeth or when teeth are in situ, because the impression material does not preserve the infinitesimal space between tightly approximated teeth.

5.3 Future work

From the findings of this study the following suggestions for future work are proposed:

- Reassessment of the phenotype of *Family I* might identify any clinical features that might be missed. With newly identified clinical assessment, another strategy of linkage analysis may be used to identify the causative genes. It may further enlighten the mode of inheritance and allows possible characterisation of disease associated variants.
- Target regions poorly covered in previous exome analysis of *Family I* to look for candidate genes that might be missed in the earlier exome sequencing analysis.
- Whole exome sequencing led to the discovery of novel mutations associated with different phenotypes of tooth agenesis. The new finding could be used to investigate the influence of those mutations on other teeth and other tissues or organs by studying their effect in transgenic mice.

- The clinical characterisation of the other 13 families could be extended to investigate the molecular basis of tooth agenesis in each family using whole exome sequencing. DNA was obtained from affected and unaffected members of those families. The discovery of novel mutations in those families could enhance knowledge about the molecular basis of tooth agenesis and lead to construction of a molecular classification of tooth agenesis.

- The initial study of tooth size in *Family I* has revealed significant reductions in tooth dimensions of several teeth. Future work on enlarged samples using the new three dimensional (3D) imaging and analysis technology may improve the study of penetration and expression of the mutations in the affected members and their relatives.

References

ABERG, T., CAVENDER, A., GAIKWAD, J. S., BRONCKERS, A. L., WANG, X., WALTIMO-SIREN, J., THESLEFF, I. & D'SOUZA, R. N. 2004. Phenotypic changes in dentition of Runx2 homozygote-null mutant mice. *J Histochem Cytochem,* 52**,** 131-9.

ADAIMY, L., CHOUERY, E., MEGARBANE, H., MROUEH, S., DELAGUE, V., NICOLAS, E., BELGUITH, H., DE MAZANCOURT, P. & MEGARBANE, A. 2007. Mutation in WNT10A is associated with an autosomal recessive ectodermal dysplasia: the odonto-onycho-dermal dysplasia. *Am J Hum Genet,* 81**,** 821-8.

AHMAD, W., BRANCOLINI, V., UL FAIYAZ, M. F., LAM, H., UL HAQUE, S., HAIDER, M., MAIMON, A., AITA, V. M., OWEN, J., BROWN, D., ZEGARELLI, D. J., AHMAD, M., OTT, J. & CHRISTIANO, A. M. 1998. A locus for autosomal recessive hypodontia with associated dental anomalies maps to chromosome 16q12.1. *Am J Hum Genet,* 62**,** 987-91.

ALAPPAT, S., ZHANG, Z. Y. & CHEN, Y. P. 2003. Msx homeobox gene family and craniofacial development. *Cell Res,* 13**,** 429-42.

AL-EMRAN, S. 1990. Prevalence of hypodontia and developmental malformation of permanent teeth in Saudi Arabian schoolchildren. *Br J Orthod,* 17**,** 115-8.

ALALUUSUA, S., CALDERARA, P., GERTHOUX, P. M., LUKINMAA, P. L., KOVERO, O., NEEDHAM, L., PATTERSON, D. G., JR., TUOMISTO, J. & MOCARELLI, P. 2004. Developmental dental aberrations after the dioxin accident in Seveso. *Environ Health Perspect,* 112**,** 1313-8.

ALVESALO, L. & PORTIN, P. 1969. The inheritance pattern of missing, peg-shaped, and strongly mesio-distally reduced upper lateral incisors. *Acta Odontol Scand,* 27**,** 563-75.

APAJALAHTI, S., ARTE, S. & PIRINEN, S. 1999. Short root anomaly in families and its association with other dental anomalies. *Eur J Oral Sci,* 107**,** 97-101.

ARTE, S., NIEMINEN, P., APAJALAHTI, S., HAAVIKKO, K., THESLEFF, I. & PIRINEN, S. 2001a. Characteristics of incisor-premolar hypodontia in families. *J Dent Res,* 80**,** 1445-50.

AVERY, J. K. & CHIEGO, D. J. 2006. *Essentials of oral histology and embryology : a clinical approach,* St. Louis, Mo. ; London, Elsevier Mosby.

AYUB, M., UR-REHMAN, F., YASINZAI, M. & AHMAD, W. 2010. A novel missense mutation in the ectodysplasin-A (EDA) gene underlies X-linked recessive nonsyndromic hypodontia. *Int J Dermatol,* 49**,** 1399-402.

BACCETTI, T. 1998a. A controlled study of associated dental anomalies. *Angle Orthod,* 68**,** 267-74.

BACCETTI, T. 1998b. Tooth rotation associated with aplasia of nonadjacent teeth. *Angle Orthod,* 68**,** 471-4.

BAILIT, H. L. 1975. Dental variation among populations. An anthropologic view. *Dent Clin North Am,* 19**,** 125-39.

BAILLEUL-FORESTIER, I., BERDAL, A., VINCKIER, F., DE RAVEL, T., FRYNS, J. P. & VERLOES, A. 2008. The genetic basis of inherited anomalies of the teeth. Part 2: syndromes with significant dental involvement. *Eur J Med Genet,* 51**,** 383-408.

BAL, E., BAALA, L., CLUZEAU, C., EL KERCH, F., OULDIM, K., HADJ-RABIA, S., BODEMER, C., MUNNICH, A., COURTOIS, G., SEFIANI, A. & SMAHI, A. 2007. Autosomal dominant anhidrotic ectodermal dysplasias at the EDARADD locus. *Hum Mutat,* 28**,** 703-9.

BAMSHAD, M. J., NG, S. B., BIGHAM, A. W., TABOR, H. K., EMOND, M. J., NICKERSON, D. A. & SHENDURE, J. 2011. Exome sequencing as a tool for Mendelian disease gene discovery. *Nat Rev Genet,* 12**,** 745-55.

BASS, T. B. 1967. Observations on the misplaced upper canine tooth. *Dent Pract Dent Rec,* 18**,** 25-33.

BAUM, B. J. & COHEN, M. M. 1971. Agenesis and tooth size in the permanent dentition. *Angle Orthod,* 41**,** 100-2.

BECKER, A., SMITH, P. & BEHAR, R. 1981. The incidence of anomalous maxillary lateral incisors in relation to palatally-displaced cuspids. *Angle Orthod,* 51**,** 24-9.

BERGENDAL, B., KLAR, J., STECKSEN-BLICKS, C., NORDERYD, J. & DAHL, N. 2011. Isolated oligodontia associated with mutations in EDARADD, AXIN2, MSX1, and PAX9 genes. *Am J Med Genet A,* 155A**,** 1616-22.

BERGSTROM, K. 1977. An orthopantomographic study of hypodontia, supernumeraries and other anomalies in school children between the ages of 8-9 years. An epidemiological study. *Swed Dent J,* 1**,** 145-57.

BJERKLIN, K., KUROL, J. & VALENTIN, J. 1992. Ectopic eruption of maxillary first permanent molars and association with other tooth and developmental disturbances. *Eur J Orthod,* 14**,** 369-75.

BLOCH-ZUPAN, A., JAMET, X., ETARD, C., LAUGEL, V., MULLER, J., GEOFFROY, V., STRAUSS, J. P., PELLETIER, V., MARION, V., POCH, O., STRAHLE, U., STOETZEL, C. & DOLLFUS, H. 2011. Homozygosity mapping and candidate prioritization identify mutations, missed by whole-exome sequencing, in SMOC2, causing major dental developmental defects. *Am J Hum Genet,* 89**,** 773-81.

BOEIRA, B. R., JR. & ECHEVERRIGRAY, S. 2013. Novel missense mutation in PAX9 gene associated with familial tooth agenesis. *J Oral Pathol Med,* 42, 99-105.

BOHRING, A., STAMM, T., SPAICH, C., HAASE, C., SPREE, K., HEHR, U., HOFFMANN, M., LEDIG, S., SEL, S., WIEACKER, P. & ROPKE, A. 2009. WNT10A mutations are a frequent cause of a broad spectrum of ectodermal dysplasias with sex-biased manifestation pattern in heterozygotes. *Am J Hum Genet,* 85**,** 97-105.

BOPP, D., BURRI, M., BAUMGARTNER, S., FRIGERIO, G. & NOLL, M. 1986. Conservation of a large protein domain in the segmentation gene paired and in functionally related genes of Drosophila. *Cell,* 47**,** 1033-40.

BORUCHOV, M. J. & GREEN, L. J. 1971. Hypodontia in human twins and families. *Am J Orthod,* 60**,** 165-74.

BRIN, I., BECKER, A. & SHALHAV, M. 1986. Position of the maxillary permanent canine in relation to anomalous or missing lateral incisors: a population study. *Eur J Orthod,* 8**,** 12-6.

BROOK, A. H. 1974. Dental anomalies of number, form and size: their prevalence in British schoolchildren. *J Int Assoc Dent Child,* 5**,** 37-53.

BROOK, A. H. 1984. A unifying aetiological explanation for anomalies of human tooth number and size. *Arch Oral Biol,* 29**,** 373-8.

BROOK, A. H., ELCOCK, C., AGGARWAL, M., LATH, D. L., RUSSELL, J. M., PATEL, P. I. & SMITH, R. N. 2009a. Tooth dimensions in hypodontia with a known PAX9 mutation. *Arch Oral Biol,* 54 Suppl 1**,** S57-62.

BROOK, A. H., GRIFFIN, R. C., SMITH, R. N., TOWNSEND, G. C., KAUR, G., DAVIS, G. R. & FEARNE, J. 2009b. Tooth size patterns in patients with hypodontia and supernumerary teeth. *Arch Oral Biol,* 54 Suppl 1**,** S63-70.

BROOK, A. H., PITTS, N. B. & RENSON, C. E. 1983. Determination of tooth dimensions from study casts using an image analysis system. *J Int Assoc Dent Child,* 14**,** 55-60.

BROOK, A. H., PITTS, N. B., YAU, F. & SANDAR, P. K. 1986. An image analysis system for the determination of tooth dimensions from study casts: comparison with manual

BUCKLEY, J. K. & DORAN, G. A. 2001. True oligodontia with microdontia in monozygous twins: implications for reappraisal of Butler's Field Theory. *Bull Group Int Rech Sci Stomatol Odontol,* 43**,** 1-10.

BUSS, P. W., HUGHES, H. E. & CLARKE, A. 1995. Twenty-four cases of the EEC syndrome: clinical presentation and management. *J Med Genet,* 32**,** 716-23.

CALLAHAN, N., MODESTO, A., MEIRA, R., SEYMEN, F., PATIR, A. & VIEIRA, A. R. 2009. Axis inhibition protein 2 (AXIN2) polymorphisms and tooth agenesis. *Arch Oral Biol,* 54**,** 45-9.

CALVANO KUCHLER, E., DE ANDRADE RISSO, P., DE CASTRO COSTA, M., MODESTO, A. & VIEIRA, A. R. 2008. Assessing the proposed association between tooth agenesis and taurodontism in 975 paediatric subjects. *Int J Paediatr Dent,* 18**,** 231-4.

CANDI, E., RUFINI, A., TERRINONI, A., GIAMBOI-MIRAGLIA, A., LENA, A. M., MANTOVANI, R., KNIGHT, R. & MELINO, G. 2007. DeltaNp63 regulates thymic development through enhanced expression of FgfR2 and Jag2. *Proc Natl Acad Sci U S A,* 104**,** 11999-2004.

CELIKOGLU, M., BAYRAM, M. & NUR, M. 2011. Patterns of third-molar agenesis and associated dental anomalies in an orthodontic population. *Am J Orthod Dentofacial Orthop,* 140**,** 856-60.

CELLI, J., DUIJF, P., HAMEL, B. C., BAMSHAD, M., KRAMER, B., SMITS, A. P., NEWBURY-ECOB, R., HENNEKAM, R. C., VAN BUGGENHOUT, G., VAN HAERINGEN, A., WOODS, C. G., VAN ESSEN, A. J., DE WAAL, R., VRIEND, G., HABER, D. A., YANG, A., MCKEON, F., BRUNNER, H. G. & VAN BOKHOVEN, H. 1999. Heterozygous germline mutations in the p53 homolog p63 are the cause of EEC syndrome. *Cell,* 99**,** 143-53.

CHEN, J., LAN, Y., BAEK, J. A., GAO, Y. & JIANG, R. 2009. Wnt/beta-catenin signaling plays an essential role in activation of odontogenic mesenchyme during early tooth development. *Dev Biol,* 334**,** 174-85.

CHEN, Y., BEI, M., WOO, I., SATOKATA, I. & MAAS, R. 1996. Msx1 controls inductive signaling in mammalian tooth morphogenesis. *Development,* 122**,** 3035-44.

CHISHTI, M. S., MUHAMMAD, D., HAIDER, M. & AHMAD, W. 2006. A novel missense mutation in MSX1 underlies autosomal recessive oligodontia with associated dental anomalies in Pakistani families. *J Hum Genet,* 51**,** 872-8.

CHOI, J. H., KIM, B. K., KIM, J. K., LEE, H. Y., PARK, J. K. & YOON, S. K. 2011. Downregulation of Foxe1 by HR suppresses Msx1 expression in the hair follicles of Hr(Hp) mice. *BMB Rep,* 44**,** 478-83.

CLAYTON, J.M. 1956. Congenital anomalies occuring in 3557 children. *J Dent Child,* 23, 206-208.

CLEMENTS, S. E., TECHANUKUL, T., LAI-CHEONG, J. E., MEE, J. B., SOUTH, A. P., POURREYRON, C., BURROWS, N. P., MELLERIO, J. E. & MCGRATH, J. A. 2012. Mutations in AEC syndrome skin reveal a role for p63 in basement membrane adhesion, skin barrier integrity and hair follicle biology. *Br J Dermatol,* 167**,** 134-44.

COLLIN, R. W., VAN DEN BORN, L. I., KLEVERING, B. J., DE CASTRO-MIRO, M., LITTINK, K. W., ARIMADYO, K., AZAM, M., YAZAR, V., ZONNEVELD, M. N., PAUN, C. C., SIEMIATKOWSKA, A. M., STROM, T. M., HEHIR-KWA, J. Y., KROES, H. Y., DE FABER, J. T., VAN SCHOONEVELD, M. J., HECKENLIVELY, J. R., HOYNG, C. B., DEN HOLLANDER, A. I. & CREMERS, F. P. 2011. High-resolution homozygosity mapping is a powerful tool to detect novel

mutations causative of autosomal recessive RP in the Dutch population. *Invest Ophthalmol Vis Sci,* 52**,** 2227-39.

COON, C. S. 1962. *The History of Man ... Second edition, revised,* pp. xxii. 438. xiv. pl. XXXII. Jonathan Cape: London; printed in U.S.A.

CORPET, F. 1988. Multiple sequence alignment with hierarchical clustering. *Nucleic Acids Res,* 16**,** 10881-90.

COUVE-PRIVAT, S., LE BRET, M., TRAIFFORT, E., QUEILLE, S., COULOMBE, J., BOUADJAR, B., AVRIL, M. F., RUAT, M., SARASIN, A. & DAYA-GROSJEAN, L. 2004. Functional analysis of novel sonic hedgehog gene mutations identified in basal cell carcinomas from xeroderma pigmentosum patients. *Cancer Res,* 64**,** 3559-65.

CSELENYI, C. S. & LEE, E. 2008. Context-dependent activation or inhibition of Wnt-beta-catenin signaling by Kremen. *Sci Signal,* 1**,** pe10.

CUDNEY, S. M. & VIEIRA, A. R. 2012. Molecular factors resulting in tooth agenesis and contemporary approaches for regeneration: a review. *Eur Arch Paediatr Dent,* 13**,** 297-304.

D'SOUZA, R. N., ABERG, T., GAIKWAD, J., CAVENDER, A., OWEN, M., KARSENTY, G. & THESLEFF, I. 1999. Cbfa1 is required for epithelial-mesenchymal interactions regulating tooth development in mice. *Development,* 126**,** 2911-20.

DAO, D. Y., YANG, X., CHEN, D., ZUSCIK, M. & O'KEEFE, R. J. 2007. Axin1 and Axin2 are regulated by TGF- and mediate cross-talk between TGF- and Wnt signaling pathways. *Ann N Y Acad Sci,* 1116**,** 82-99.

DAO, D. Y., YANG, X., FLICK, L. M., CHEN, D., HILTON, M. J. & O'KEEFE, R. J. 2010. Axin2 regulates chondrocyte maturation and axial skeletal development. *J Orthop Res,* 28**,** 89-95.

DAS, P., HAI, M., ELCOCK, C., LEAL, S. M., BROWN, D. T., BROOK, A. H. & PATEL, P. I. 2003. Novel missense mutations and a 288-bp exonic insertion in PAX9 in families with autosomal dominant hypodontia. *Am J Med Genet A,* 118A**,** 35-42.

DASSULE, H. R., LEWIS, P., BEI, M., MAAS, R. & MCMAHON, A. P. 2000. Sonic hedgehog regulates growth and morphogenesis of the tooth. *Development,* 127**,** 4775-85.

DASSULE, H. R. & MCMAHON, A. P. 1998. Analysis of epithelial-mesenchymal interactions in the initial morphogenesis of the mammalian tooth. *Dev Biol,* 202**,** 215-27.

DAUGAARD-JENSEN, J., NODAL, M. & KJAER, I. 1997a. Pattern of agenesis in the primary dentition: a radiographic study of 193 cases. *Int J Paediatr Dent,* 7**,** 3-7.

DAUGAARD-JENSEN, J., NODAL, M., SKOVGAARD, L. T. & KJAER, I. 1997b. Comparison of the pattern of agenesis in the primary and permanent dentitions in a population characterized by agenesis in the primary dentition. *Int J Paediatr Dent,* 7**,** 143-8.

DAVIS, P. J. 1987. Hypodontia and hyperdontia of permanent teeth in Hong Kong schoolchildren. *Community Dent Oral Epidemiol,* 15**,** 218-20.

DE MUYNCK, S., SCHOLLEN, E., MATTHIJS, G., VERDONCK, A., DEVRIENDT, K. & CARELS, C. 2004. A novel MSX1 mutation in hypodontia. *Am J Med Genet A,* 128A**,** 401-3.

DEN DUNNEN, J. T. & ANTONARAKIS, S. E. 2001. Nomenclature for the description of human sequence variations. *Hum Genet,* 109**,** 121-4.

DESMYTER, L., GHASSIBE, M., REVENCU, N., BOUTE, O., LEES, M., FRANCOIS, G., VERELLEN-DUMOULIN, C., SZNAJER, Y., MONCLA, A., BENATEAU, H., CLAES, K., DEVRIENDT, K., MATHIEU, M., VAN MALDERGEM, L., ADDOR,

M. C., DROUIN-GARRAUD, V., MORTIER, G., BOUMA, M., DIEUX-COESLIER, A., GENEVIEVE, D., GOLDENBERG, A., GOZU, A., MAKRYTHANASIS, P., MCENTAGART, U., SANCHEZ, A., VILAIN, C., VERMEER, S., CONNELL, F., VERHEIJ, J., MANOUVRIER, S., PIERQUIN, G., ODENT, S., HOLDER-ESPINASSE, M., VINCENT-DELORME, C., GILLEROT, Y., VANWIJCK, R., BAYET, B. & VIKKULA, M. 2010. IRF6 Screening of Syndromic and a priori Non-Syndromic Cleft Lip and Palate Patients: Identification of a New Type of Minor VWS Sign. *Mol Syndromol,* 1**,** 67-74.

DIAMOND, E., AMEN, M., HU, Q., ESPINOZA, H. M. & AMENDT, B. A. 2006. Functional interactions between Dlx2 and lymphoid enhancer factor regulate Msx2. *Nucleic Acids Res,* 34**,** 5951-65.

DOOLITTLE, W. F. 2013. Is junk DNA bunk? A critique of ENCODE. *Proc Natl Acad Sci U S A.*

DRESSLER, S., MEYER-MARCOTTY, P., WEISSCHUH, N., JABLONSKI-MOMENI, A., PIEPER, K., GRAMER, G. & GRAMER, E. 2010. Dental and Craniofacial Anomalies Associated with Axenfeld-Rieger Syndrome with PITX2 Mutation. *Case Report Med,* 2010**,** 621984.

DUNHAM, I., KUNDAJE, A., ALDRED, S. F., COLLINS, P. J., DAVIS, C. A., DOYLE, F., EPSTEIN, C. B., FRIETZE, S., HARROW, J., KAUL, R., KHATUN, J., LAJOIE, B. R., LANDT, S. G., LEE, B. K., PAULI, F., ROSENBLOOM, K. R., SABO, P., SAFI, A., SANYAL, A., SHORESH, N., SIMON, J. M., SONG, L., TRINKLEIN, N. D., ALTSHULER, R. C., BIRNEY, E., BROWN, J. B., CHENG, C., DJEBALI, S., DONG, X., ERNST, J., FUREY, T. S., GERSTEIN, M., GIARDINE, B., GREVEN, M., HARDISON, R. C., HARRIS, R. S., HERRERO, J., HOFFMAN, M. M., IYER, S., KELLLIS, M., KHERADPOUR, P., LASSMAN, T., LI, Q., LIN, X., MARINOV, G. K., MERKEL, A., MORTAZAVI, A., PARKER, S. C., REDDY, T. E., ROZOWSKY, J., SCHLESINGER, F., THURMAN, R. E., WANG, J., WARD, L. D., WHITFIELD, T. W., WILDER, S. P., WU, W., XI, H. S., YIP, K. Y., ZHUANG, J., BERNSTEIN, B. E., GREEN, E. D., GUNTER, C., SNYDER, M., PAZIN, M. J., LOWDON, R. F., DILLON, L. A., ADAMS, L. B., KELLY, C. J., ZHANG, J., WEXLER, J. R., GOOD, P. J., FEINGOLD, E. A., CRAWFORD, G. E., DEKKER, J., ELINITSKI, L., FARNHAM, P. J., GIDDINGS, M. C., GINGERAS, T. R., GUIGO, R., HUBBARD, T. J., KELLIS, M., KENT, W. J., LIEB, J. D., MARGULIES, E. H., MYERS, R. M., STARNATOYANNOPOULOS, J. A., TENNEBAUM, S. A., WENG, Z., WHITE, K. P., WOLD, B., YU, Y., WROBEL, J., RISK, B. A., GUNAWARDENA, H. P., KUIPER, H. C., MAIER, C. W., XIE, L., CHEN, X., MIKKELSEN, T. S., *et al*. 2012. An integrated encyclopedia of DNA elements in the human genome. *Nature,* 489**,** 57-74.

DWYER, E., HYLAND, J., MODAFF, P. & PAULI, R. M. 2010. Genotype-phenotype correlation in DTDST dysplasias: Atelosteogenesis type II and diastrophic dysplasia variant in one family. *Am J Med Genet A,* 152A**,** 3043-50.

EDWARDS, D. M. & ROBERTS, G. J. 1990. Short root anomaly. *Br Dent J,* 169**,** 292-3.

EGERMARK-ERIKSSON, I. & LIND, V. 1971. Congenital numerical variation in the permanent dentition. D. Sex distribution of hypodontia and hyperodontia. *Odontol Revy,* 22**,** 309-15.

EICHBERGER, T., REGL, G., IKRAM, M. S., NEILL, G. W., PHILPOTT, M. P., ABERGER, F. & FRISCHAUF, A. M. 2004. FOXE1, a new transcriptional target of GLI2 is expressed in human epidermis and basal cell carcinoma. *J Invest Dermatol,* 122**,** 1180-7.

ERICSON, S. & KUROL, J. 1986. Radiographic assessment of maxillary canine eruption in children with clinical signs of eruption disturbance. *Eur J Orthod,* 8**,** 133-40.

FERRANTE, M. I., GIORGIO, G., FEATHER, S. A., BULFONE, A., WRIGHT, V., GHIANI, M., SELICORNI, A., GAMMARO, L., SCOLARI, F., WOOLF, A. S., SYLVIE, O., BERNARD, L., MALCOLM, S., WINTER, R., BALLABIO, A. & FRANCO, B. 2001. Identification of the gene for oral-facial-digital type I syndrome. *Am J Hum Genet,* 68**,** 569-76.

FRAZIER-BOWERS, S. A., GUO, D. C., CAVENDER, A., XUE, L., EVANS, B., KING, T., MILEWICZ, D. & D'SOUZA, R. N. 2002. A novel mutation in human PAX9 causes molar oligodontia. *J Dent Res,* 81**,** 129-33.

FREIRE-MAIA, N. 1971. Ectodermal dysplasias. *Hum Hered,* 21**,** 309-12.

FREIRE-MAIA, N., LISBOA-COSTA, T. & PAGNAN, N. A. 2001. Ectodermal dysplasias: how many? *Am J Med Genet,* 104**,** 84.

GALLUCCIO, G., CASTELLANO, M. & LA MONACA, C. 2012. Genetic basis of non-syndromic anomalies of human tooth number. *Arch Oral Biol,* 57**,** 918-30.

GAMAZON, E. R., ZHANG, W., DOLAN, M. E. & COX, N. J. 2010. Comprehensive survey of SNPs in the Affymetrix exon array using the 1000 Genomes dataset. *PLoS One,* 5**,** e9366.

GAO, C. & CHEN, Y. G. 2010. Dishevelled: The hub of Wnt signaling. *Cell Signal,* 22**,** 717-27.

GARIB, D. G., ALENCAR, B. M., LAURIS, J. R. & BACCETTI, T. 2010. Agenesis of maxillary lateral incisors and associated dental anomalies. *Am J Orthod Dentofacial Orthop,* 137**,** 732 e1-6; discussion 732-3.

GARIB, D. G., PECK, S. & GOMES, S. C. 2009. Increased occurrence of dental anomalies associated with second-premolar agenesis. *Angle Orthod,* 79**,** 436-41.

GARN, S. M. & LEWIS, A. B. 1970. The gradient and the pattern of crown-size reduction in simple hypodontia. *Angle Orthod,* 40**,** 51-8.

GARN, S. M., LEWIS, A. B. & BONNE, B. 1961. Third molar polymorphism and the timing of tooth formation. *Nature,* 192**,** 989.

GASS, J. K., WILSON, N. J., SMITH, F. J., LANE, E. B., MCLEAN, W. H., RYTINA, E., SALVARY, I. & BURROWS, N. P. 2009. Steatocystoma multiplex, oligodontia and partial persistent primary dentition associated with a novel keratin 17 mutation. *Br J Dermatol,* 161**,** 1396-8.

GOMES, R. R., HABCKOST, C. D., JUNQUEIRA, L. G., LEITE, A. F., FIGUEIREDO, P. T., PAULA, L. M. & ACEVEDO, A. C. 2012. Taurodontism in Brazilian patients with tooth agenesis and first and second-degree relatives: a case-control study. *Arch Oral Biol,* 57**,** 1062-9.

GOYA, H. A., TANAKA, S., MAEDA, T. & AKIMOTO, Y. 2008. An orthopantomographic study of hypodontia in permanent teeth of Japanese pediatric patients. *J Oral Sci,* 50**,** 143-50.

GRAHNEN , H.1956. Hypodontia in the permanent dentition. *Odontol Revy,* 7,1–100

GRAHNEN, H, GRANATH, L.E. 1961. Numerical variations in primary dentition and their correlation with the permanent dentition. *Odont Revoy,* 12, 348-57.

GRIGORIOU, M., TUCKER, A. S., SHARPE, P. T. & PACHNIS, V. 1998. Expression and regulation of Lhx6 and Lhx7, a novel subfamily of LIM homeodomain encoding genes, suggests a role in mammalian head development. *Development,* 125**,** 2063-74.

GRITLI-LINDE, A., BEI, M., MAAS, R., ZHANG, X. M., LINDE, A. & MCMAHON, A. P. 2002. Shh signaling within the dental epithelium is necessary for cell proliferation, growth and polarization. *Development,* 129**,** 5323-37.

GROVER, P. S. & LORTON, L. 1985. The incidence of unerupted permanent teeth and related clinical cases. *Oral Surg Oral Med Oral Pathol,* 59**,** 420-5.

HAAVIKKO, K. 1971. Hypodontia of permanent teeth. An orthopantomographic study. *Suom Hammaslaak Toim,* 67**,** 219-25.

HAGHIGHI, A., NIKUEI, P., HAGHIGHI-KAKHKI, H., N, H. S.-G., BAGHESTANI, S., KRAWITZ, P. M., HECHT, J. & MUNDLOS, S. 2012. Identification of a Novel Missense Mutation in EDAR Causing Autosomal Recessive Hypohidrotic Ectodermal Dysplasia with Bilateral Amastia and Palmoplantar Hyperkeratosis. *Br J Dermatol*.

HAMADA, H., MENO, C., WATANABE, D. & SAIJOH, Y. 2002. Establishment of vertebrate left-right asymmetry. *Nat Rev Genet,* 3**,** 103-13.

HAN, D., GONG, Y., WU, H., ZHANG, X., YAN, M., WANG, X., QU, H., FENG, H. & SONG, S. 2008. Novel EDA mutation resulting in X-linked non-syndromic hypodontia and the pattern of EDA-associated isolated tooth agenesis. *Eur J Med Genet,* 51**,** 536-46.

HANSEN, L., RIIS, A. K., SILAHTAROGLU, A., HOVE, H., LAURIDSEN, E., EIBERG, H. & KREIBORG, S. 2011. RUNX2 analysis of Danish cleidocranial dysplasia families. *Clin Genet,* 79**,** 254-63.

HARDCASTLE, Z., MO, R., HUI, C. C. & SHARPE, P. T. 1998. The Shh signalling pathway in tooth development: defects in Gli2 and Gli3 mutants. *Development,* 125**,** 2803-11.

HARJUNMAA, E., KALLONEN, A., VOUTILAINEN, M., HAMALAINEN, K., MIKKOLA, M. L. & JERNVALL, J. 2012. On the difficulty of increasing dental complexity. *Nature,* 483**,** 324-7.

HEADON, D. J., EMMAL, S. A., FERGUSON, B. M., TUCKER, A. S., JUSTICE, M. J., SHARPE, P. T., ZONANA, J. & OVERBEEK, P. A. 2001. Gene defect in ectodermal dysplasia implicates a death domain adapter in development. *Nature,* 414**,** 913-6.

HIROKAWA, N., TANAKA, Y. & OKADA, Y. 2009. Left-right determination: involvement of molecular motor KIF3, cilia, and nodal flow. *Cold Spring Harb Perspect Biol,* 1**,** a000802.

HOLTTA, P., ALALUUSUA, S., SAARINEN-PIHKALA, U. M., PELTOLA, J. & HOVI, L. 2005a. Agenesis and microdontia of permanent teeth as late adverse effects after stem cell transplantation in young children. *Cancer,* 103**,** 181-90.

HOLTTA, P., HOVI, L., SAARINEN-PIHKALA, U. M., PELTOLA, J. & ALALUUSUA, S. 2005b. Disturbed root development of permanent teeth after pediatric stem cell transplantation. Dental root development after SCT. *Cancer,* 103**,** 1484-93.

HOLTTA, P., KIVIRANTA, H., LEPPANIEMI, A., VARTIAINEN, T., LUKINMAA, P. L. & ALALUUSUA, S. 2001. Developmental dental defects in children who reside by a river polluted by dioxins and furans. *Arch Environ Health,* 56**,** 522-8.

HOSUR, M. B., PURANIC, R S., and VANAKI, S. 2011. Oligodontia: A casereport and Review of Litreature. *World J Dent,* 2, 259-262.

IRVINE, A. D. 2009. Towards a unified classification of the ectodermal dysplasias: opportunities outweigh challenges. *Am J Med Genet A,* 149A**,** 1970-2.

ITIN, P. H. 2009. Rationale and background as basis for a new classification of the ectodermal dysplasias. *Am J Med Genet A,* 149A**,** 1973-6.

ISRAELI, S., GOLDBERG, I., FUCHS-TELEM, D., BERGMAN, R., INDELMAN, M., BITTERMAN-DEUTSCH, O., HAREL, A., MASHIACH, Y., SARIG, O. & SPRECHER, E. 2013. Non-syndromic autosomal recessive congenital ichthyosis in the Israeli population. *Clin Exp Dermatol*.

JARVINEN, S. & LEHTINEN, L. 1981. Supernumerary and congenitally missing primary teeth in Finnish children. An epidemiologic study. *Acta Odontol Scand,* 39**,** 83-6.

JEZEWSKI, P. A., VIEIRA, A. R., NISHIMURA, C., LUDWIG, B., JOHNSON, M., O'BRIEN, S. E., DAACK-HIRSCH, S., SCHULTZ, R. E., WEBER, A., NEPOMUCENA, B., ROMITTI, P. A., CHRISTENSEN, K., ORIOLI, I. M., CASTILLA, E. E., MACHIDA, J., NATSUME, N. & MURRAY, J. C. 2003. Complete sequencing shows a role for MSX1 in non-syndromic cleft lip and palate. *J Med Genet,* 40**,** 399-407.

JORGENSON, R. J. 1980. Clinician's view of hypodontia. *J Am Dent Assoc,* 101**,** 283-6.

JUMLONGRAS, D., BEI, M., STIMSON, J. M., WANG, W. F., DEPALMA, S. R., SEIDMAN, C. E., FELBOR, U., MAAS, R., SEIDMAN, J. G. & OLSEN, B. R. 2001. A nonsense mutation in MSX1 causes Witkop syndrome. *Am J Hum Genet,* 69**,** 67-74.

JUMLONGRAS, D., LIN, J. Y., CHAPRA, A., SEIDMAN, C. E., SEIDMAN, J. G., MAAS, R. L. & OLSEN, B. R. 2004. A novel missense mutation in the paired domain of PAX9 causes non-syndromic oligodontia. *Hum Genet,* 114**,** 242-9.

KANTAPUTRA, P. & SRIPATHOMSAWAT, W. 2011. WNT10A and isolated hypodontia. *Am J Med Genet A,* 155A**,** 1119-22.

KAPADIA, H., FRAZIER-BOWERS, S., OGAWA, T. & D'SOUZA, R. N. 2006. Molecular characterization of a novel PAX9 missense mutation causing posterior tooth agenesis. *Eur J Hum Genet,* 14**,** 403-9.

KERE, J., SRIVASTAVA, A. K., MONTONEN, O., ZONANA, J., THOMAS, N., FERGUSON, B., MUNOZ, F., MORGAN, D., CLARKE, A., BAYBAYAN, P., CHEN, E. Y., EZER, S., SAARIALHO-KERE, U., DE LA CHAPELLE, A. & SCHLESSINGER, D. 1996. X-linked anhidrotic (hypohidrotic) ectodermal dysplasia is caused by mutation in a novel transmembrane protein. *Nat Genet,* 13**,** 409-16.

KHATUN, J., YU, Y., WROBEL, J. A., RISK, B. A., GUNAWARDENA, H. P., SECREST, A., SPITZER, W. J., XIE, L., WANG, L., CHEN, X. & GIDDINGS, M. C. 2013. Whole human genome proteogenomic mapping for ENCODE cell line data: identifying protein-coding regions. *BMC Genomics,* 14**,** 141.

KIM, J. W., SIMMER, J. P., LIN, B. P. & HU, J. C. 2006. Novel MSX1 frameshift causes autosomal-dominant oligodontia. *J Dent Res,* 85**,** 267-71.

KIST, R., WATSON, M., WANG, X., CAIRNS, P., MILES, C., REID, D. J. & PETERS, H. 2005. Reduction of Pax9 gene dosage in an allelic series of mouse mutants causes hypodontia and oligodontia. *Hum Mol Genet,* 14**,** 3605-17.

KJAER, I. 1997. Can the location of tooth agenesis and the location of initial bone loss seen in juvenile periodontitis be explained by neural developmental fields in the jaws? *Acta Odontol Scand,* 55**,** 70-2.

KLEIN, M. L., NIEMINEN, P., LAMMI, L., NIEBUHR, E. & KREIBORG, S. 2005. Novel mutation of the initiation codon of PAX9 causes oligodontia. *J Dent Res,* 84**,** 43-7.

KLINGENBERG, C. P., BARLUENGA, M. & MEYER, A. 2002. Shape analysis of symmetric structures: quantifying variation among individuals and asymmetry. *Evolution,* 56**,** 1909-20.

KOLLAR, E. J. & BAIRD, G. R. 1969. The influence of the dental papilla on the development of tooth shape in embryonic mouse tooth germs. *J Embryol Exp Morphol,* 21**,** 131-48.

KOLLAR, E. J. & BAIRD, G. R. 1970. Tissue interactions in embryonic mouse tooth germs. I. Reorganization of the dental epithelium during tooth-germ reconstruction. *J Embryol Exp Morphol,* 24**,** 159-71.

KRUGLYAK, L., DALY, M. J. & LANDER, E. S. 1995. Rapid multipoint linkage analysis of recessive traits in nuclear families, including homozygosity mapping. *Am J Hum Genet,* 56**,** 519-27.

KU, C. S., COOPER, D. N., POLYCHRONAKOS, C., NAIDOO, N., WU, M. & SOONG, R. 2012. Exome sequencing: dual role as a discovery and diagnostic tool. *Ann Neurol,* 71**,** 5-14.

KU, C. S., NAIDOO, N. & PAWITAN, Y. 2011. Revisiting Mendelian disorders through exome sequencing. *Hum Genet,* 129**,** 351-70.

LAI, P. Y. & SEOW, W. K. 1989. A controlled study of the association of various dental anomalies with hypodontia of permanent teeth. *Pediatr Dent,* 11**,** 291-6.

LAMMI, L., ARTE, S., SOMER, M., JARVINEN, H., LAHERMO, P., THESLEFF, I., PIRINEN, S. & NIEMINEN, P. 2004a. Mutations in AXIN2 cause familial tooth agenesis and predispose to colorectal cancer. *Am J Hum Genet,* 74**,** 1043-50.

LAMMI, L., ARTE, S., SOMER, M., JARVINEN, H., LAHERMO, P., THESLEFF, I., PIRINEN, S. & NIEMINEN, P. 2004b. Mutations in AXIN2 Cause Familial Tooth Agenesis and Predispose to Colorectal Cancer. *The American Journal of Human Genetics,* 74**,** 1043-1050.

LAMMI, L., HALONEN, K., PIRINEN, S., THESLEFF, I., ARTE, S. & NIEMINEN, P. 2003. A missense mutation in PAX9 in a family with distinct phenotype of oligodontia. *Eur J Hum Genet,* 11**,** 866-71.

LANDER, E. S. 2011. Initial impact of the sequencing of the human genome. *Nature,* 470**,** 187-97.

LARKIN, M. A., BLACKSHIELDS, G., BROWN, N. P., CHENNA, R., MCGETTIGAN, P. A., MCWILLIAM, H., VALENTIN, F., WALLACE, I. M., WILM, A., LOPEZ, R., THOMPSON, J. D., GIBSON, T. J. & HIGGINS, D. G. 2007. Clustal W and Clustal X version 2.0. *Bioinformatics,* 23**,** 2947-8.

LAURIKKALA, J., MIKKOLA, M. L., JAMES, M., TUMMERS, M., MILLS, A. A. & THESLEFF, I. 2006. p63 regulates multiple signalling pathways required for ectodermal organogenesis and differentiation. *Development,* 133**,** 1553-63.

LEKKAS, C., LATIEF, B. S., TER RAHE, S. P. & KUIJPERS-JAGTMAN, A. M. 2000. The adult unoperated cleft patient: absence of maxillary teeth outside the cleft area. *Cleft Palate Craniofac J,* 37**,** 17-20.

LI, S., LI, J., CHENG, J., ZHOU, B., TONG, X., DONG, X., WANG, Z., HU, Q., CHEN, M. & HUA, Z. C. 2008. Non-syndromic tooth agenesis in two Chinese families associated with novel missense mutations in the TNF domain of EDA (ectodysplasin A). *PLoS One,* 3**,** e2396.

LIDRAL, A. C. & REISING, B. C. 2002. The role of MSX1 in human tooth agenesis. *J Dent Res,* 81**,** 274-8.

LIN, C. R., KIOUSSI, C., O'CONNELL, S., BRIATA, P., SZETO, D., LIU, F., IZPISUA-BELMONTE, J. C. & ROSENFELD, M. G. 1999. Pitx2 regulates lung asymmetry, cardiac positioning and pituitary and tooth morphogenesis. *Nature,* 401**,** 279-82.

LIND, V. 1972. Short root anomaly. *Scand J Dent Res,* 80**,** 85-93.

LIU, P., LU, J., CARDOSO, W. V. & VAZIRI, C. 2008. The SPARC-related factor SMOC-2 promotes growth factor-induced cyclin D1 expression and DNA synthesis via integrin-linked kinase. *Mol Biol Cell,* 19**,** 248-61.

LIU, W., WANG, H., ZHAO, S., ZHAO, W., BAI, S., ZHAO, Y., XU, S., WU, C., HUANG, W., CHEN, Z., FENG, G. & HE, L. 2001. The novel gene locus for agenesis of permanent teeth (He-Zhao deficiency) maps to chromosome 10q11.2. *J Dent Res,* 80**,** 1716-20.

LOURENCO RIBEIRO, L., TEIXEIRA DAS NEVES, L., COSTA, B. & RIBEIRO GOMIDE, M. 2003. Dental anomalies of the permanent lateral incisors and prevalence of hypodontia outside the cleft area in complete unilateral cleft lip and palate. *Cleft Palate Craniofac J,* 40**,** 172-5.

LU, M. F., PRESSMAN, C., DYER, R., JOHNSON, R. L. & MARTIN, J. F. 1999. Function of Rieger syndrome gene in left-right asymmetry and craniofacial development. *Nature,* 401**,** 276-8.

LU, P. D. & SCHAFFER, J. V. 2008. Hypohidrotic ectodermal dysplasia. *Dermatol Online J,* 14**,** 22.

LUKINMAA, P. L., SAHLBERG, C., LEPPANIEMI, A., PARTANEN, A. M., KOVERO, O., POHJANVIRTA, R., TUOMISTO, J. & ALALUUSUA, S. 2001. Arrest of rat molar tooth development by lactational exposure to 2,3,7,8-tetrachlorodibenzo-p-dioxin. *Toxicol Appl Pharmacol,* 173**,** 38-47.

LUMSDEN, A. G. 1988. Spatial organization of the epithelium and the role of neural crest cells in the initiation of the mammalian tooth germ. *Development,* 103 Suppl**,** 155-69.

LUNDSTORM, A1960., Asymmetries in the number and size of the teeth and thier aetiological significance. Transactions. *Euro Ortho Soc,* 1, 167-185.

LUSSI, A. & JAEGGI, T. 2008. Erosion--diagnosis and risk factors. *Clin Oral Investig,* 12 Suppl 1**,** S5-13.

MAGNUSSON, T. E. 1977. Prevalence of hypodontia and malformations of permanent teeth in Iceland. *Community Dent Oral Epidemiol,* 5**,** 173-8.

MAJEWSKI, J., SCHWARTZENTRUBER, J., LALONDE, E., MONTPETIT, A. & JABADO, N. 2011. What can exome sequencing do for you? *J Med Genet,* 48**,** 580-9.

MARIAN, A. J. 2012a. Challenges in Medical Applications of Whole Exome/Genome Sequencing Discoveries. *Trends Cardiovasc Med*.

MARIAN, A. J. 2012b. Molecular genetic studies of complex phenotypes. *Transl Res,* 159**,** 64-79.

MATALOVA, E., FLEISCHMANNOVA, J., SHARPE, P. T. & TUCKER, A. S. 2008. Tooth agenesis: from molecular genetics to molecular dentistry. *J Dent Res,* 87**,** 617-23.

MATTHEEUWS, N., DERMAUT, L. & MARTENS, G. 2004. Has hypodontia increased in Caucasians during the 20th century? A meta-analysis. *Eur J Orthod,* 26**,** 99-103.

MCGRATH, J. A., DUIJF, P. H., DOETSCH, V., IRVINE, A. D., DE WAAL, R., VANMOLKOT, K. R., WESSAGOWIT, V., KELLY, A., ATHERTON, D. J., GRIFFITHS, W. A., ORLOW, S. J., VAN HAERINGEN, A., AUSEMS, M. G., YANG, A., MCKEON, F., BAMSHAD, M. A., BRUNNER, H. G., HAMEL, B. C. & VAN BOKHOVEN, H. 2001. Hay-Wells syndrome is caused by heterozygous missense mutations in the SAM domain of p63. *Hum Mol Genet,* 10**,** 221-9.

MCKEOWN, H. F., ROBINSON, D. L., ELCOCK, C., AL-SHAROOD, M. & BROOK, A. H. 2002. Tooth dimensions in hypodontia patients, their unaffected relatives and a control group measured by a new image analysis system. *Eur J Orthod,* 24**,** 131-41.

MILLS, A. A., ZHENG, B., WANG, X. J., VOGEL, H., ROOP, D. R. & BRADLEY, A. 1999. p63 is a p53 homologue required for limb and epidermal morphogenesis. *Nature,* 398**,** 708-13.

MINA, M. & KOLLAR, E. J. 1987. The induction of odontogenesis in non-dental mesenchyme combined with early murine mandibular arch epithelium. *Arch Oral Biol,* 32**,** 123-7.

MITSIADIS, T. A., ANGELI, I., JAMES, C., LENDAHL, U. & SHARPE, P. T. 2003. Role of Islet1 in the patterning of murine dentition. *Development,* 130**,** 4451-60.

MITSIADIS, T. A. & SMITH, M. M. 2006. How do genes make teeth to order through development? *J Exp Zool B Mol Dev Evol,* 306**,** 177-82.

MOLLER, P., BERG, K., RUUD, A. F. & KVIEN, T. K. 1981. Variable expression of familial hypodontia in monozygotic triplets. *Scand J Dent Res,* 89**,** 16-8.

MONREAL, A. W., ZONANA, J. & FERGUSON, B. 1998. Identification of a new splice form of the EDA1 gene permits detection of nearly all X-linked hypohidrotic ectodermal dysplasia mutations. *Am J Hum Genet,* 63**,** 380-9.

MOORREES, C. F. & REED, R. B. 1964. Correlations among Crown Diameters of Human Teeth. *Arch Oral Biol,* 9**,** 685-97.

MOSTOWSKA, A., BIEDZIAK, B. & JAGODZINSKI, P. P. 2006a. Axis inhibition protein 2 (AXIN2) polymorphisms may be a risk factor for selective tooth agenesis. *J Hum Genet,* 51**,** 262-6.

MOSTOWSKA, A., BIEDZIAK, B. & JAGODZINSKI, P. P. 2012. Novel MSX1 mutation in a family with autosomal-dominant hypodontia of second premolars and third molars. *Arch Oral Biol,* 57**,** 790-5.

MOSTOWSKA, A., BIEDZIAK, B. & TRZECIAK, W. H. 2006b. A novel c.581C>T transition localized in a highly conserved homeobox sequence of MSX1: is it responsible for oligodontia? *J Appl Genet,* 47**,** 159-64.

MOSTOWSKA, A., KOBIELAK, A., BIEDZIAK, B. & TRZECIAK, W. H. 2003. Novel mutation in the paired box sequence of PAX9 gene in a sporadic form of oligodontia. *Eur J Oral Sci,* 111**,** 272-6.

MULLER, T. P., HILL, I. N., PETERSON, A. C. & BLAYNEY, J. R. 1970. A survey of congenitally missing permanent teeth. *J Am Dent Assoc,* 81**,** 101-7.

MUNDLOS, S., OTTO, F., MUNDLOS, C., MULLIKEN, J. B., AYLSWORTH, A. S., ALBRIGHT, S., LINDHOUT, D., COLE, W. G., HENN, W., KNOLL, J. H., OWEN, M. J., MERTELSMANN, R., ZABEL, B. U. & OLSEN, B. R. 1997. Mutations involving the transcription factor CBFA1 cause cleidocranial dysplasia. *Cell,* 89**,** 773-9.

MURRAY, B. 2012. Arrhythmogenic right ventricular dysplasia/cardiomyopathy (ARVD/C): a review of molecular and clinical literature. *J Genet Couns,* 21**,** 494-504.

NANCI, A. & TEN CATE, A. R. O. H. 2008. *Ten Cate's oral histology : development, structure, and function,* St. Louis, Mo. ; [London], Mosby.

NAQVI, S. K., WASIF, N., JAVAID, H. & AHMAD, W. 2011. Two novel mutations in the gene EDAR causing autosomal recessive hypohidrotic ectodermal dysplasia. *Orthod Craniofac Res,* 14**,** 156-9.

NAWAZ, S., KLAR, J., WAJID, M., ASLAM, M., TARIQ, M., SCHUSTER, J., BAIG, S. M. & DAHL, N. 2009. WNT10A missense mutation associated with a complete odonto-onycho-dermal dysplasia syndrome. *Eur J Hum Genet,* 17**,** 1600-5.

NIEMINEN, P. 2009. Genetic basis of tooth agenesis. *J Exp Zool B Mol Dev Evol,* 312B**,** 320-42.

NIEMINEN, P., ARTE, S., TANNER, D., PAULIN, L., ALALUUSUA, S., THESLEFF, I. & PIRINEN, S. 2001. Identification of a nonsense mutation in the PAX9 gene in molar oligodontia. *Eur J Hum Genet,* 9**,** 743-6.

NIEMINEN, P., KOTILAINEN, J., AALTO, Y., KNUUTILA, S., PIRINEN, S. & THESLEFF, I. 2003. MSX1 gene is deleted in Wolf-Hirschhorn syndrome patients with oligodontia. *J Dent Res,* 82**,** 1013-7.

NISWANDER, J. D. & SUJAKU, C. 1963. Congenital Anomalies of Teeth in Japanese Children. *Am J Phys Anthropol,* 21**,** 569-74.

OLIVER, R. G., MANNION, J. E. & ROBINSON, J. M. 1989. Morphology of the maxillary lateral incisor in cases of unilateral impaction of the maxillary canine. *Br J Orthod,* 16**,** 9-16.

PANGANIBAN, G. & RUBENSTEIN, J. L. 2002. Developmental functions of the Distal-less/Dlx homeobox genes. *Development,* 129**,** 4371-86.

PAWLOWSKA, E., JANIK-PAPIS, K., WISNIEWSKA-JAROSINSKA, M., SZCZEPANSKA, J. & BLASIAK, J. 2009. Mutations in the human homeobox MSX1 gene in the congenital lack of permanent teeth. *Tohoku J Exp Med,* 217**,** 307-12.

PAZIN, D. E. & ALBRECHT, K. H. 2009. Developmental expression of Smoc1 and Smoc2 suggests potential roles in fetal gonad and reproductive tract differentiation. *Dev Dyn,* 238**,** 2877-90.

PECK, S., PECK, L. & KATAJA, M. 1994. The palatally displaced canine as a dental anomaly of genetic origin. *Angle Orthod,* 64**,** 249-56.

PEDERSEN, L. B., CLAUSEN, N., SCHRODER, H., SCHMIDT, M. & POULSEN, S. 2012. Microdontia and hypodontia of premolars and permanent molars in childhood cancer survivors after chemotherapy. *Int J Paediatr Dent,* 22**,** 239-43.

PETERS, H., NEUBUSER, A., KRATOCHWIL, K. & BALLING, R. 1998. Pax9-deficient mice lack pharyngeal pouch derivatives and teeth and exhibit craniofacial and limb abnormalities. *Genes Dev,* 12**,** 2735-47.

PINDBORG, J. J. 1970. *Pathology of dental hard tissues,* Copenhagen, Munksgaard.

PINHEIRO, M. & FREIRE-MAIA, N. 1994. Ectodermal dysplasias: a clinical classification and a causal review. *Am J Med Genet,* 53**,** 153-62.

PIO, S. F., OLIVEIRA, G. C. & REZENDE, S. M. 2009. [The molecular bases of hemophilia A.]. *Rev Assoc Med Bras,* 55**,** 213-9.

PIRINEN, S., ARTE, S. & APAJALAHTI, S. 1996. Palatal displacement of canine is genetic and related to congenital absence of teeth. *J Dent Res,* 75**,** 1742-6.

PIRINEN, S., KENTALA, A., NIEMINEN, P., VARILO, T., THESLEFF, I. & ARTE, S. 2001. Recessively inherited lower incisor hypodontia. *J Med Genet,* 38**,** 551-6.

PISPA, J. & THESLEFF, I. 2003. Mechanisms of ectodermal organogenesis. *Dev Biol,* 262**,** 195-205.

POLDER, B. J., VAN'T HOF, M. A., VAN DER LINDEN, F. P. & KUIJPERS-JAGTMAN, A. M. 2004. A meta-analysis of the prevalence of dental agenesis of permanent teeth. *Community Dent Oral Epidemiol,* 32**,** 217-26.

PRIOLO, M. & LAGANA, C. 2001. Ectodermal dysplasias: a new clinical-genetic classification. *J Med Genet,* 38**,** 579-85.

RANTA, R. 1986. A review of tooth formation in children with cleft lip/palate. *Am J Orthod Dentofacial Orthop,* 90**,** 11-8.

RANTA, R. & TULENSALO, T. 1988. Symmetry and combinations of hypodontia in non-cleft and cleft palate children. *Scand J Dent Res,* 96**,** 1-8.

RASOOL, M., SCHUSTER, J., ASLAM, M., TARIQ, M., AHMAD, I., ALI, A., ENTESARIAN, M., DAHL, N. & BAIG, S. M. 2008. A novel missense mutation in the EDA gene associated with X-linked recessive isolated hypodontia. *J Hum Genet,* 53**,** 894-8.

RAVN, J. J. 1971. Aplasia, supernumerary teeth and fused teeth in the primary dentition. An epidemiologic study. *Scand J Dent Res,* 79**,** 1-6.

RITCHIE, M. E., LIU, R., CARVALHO, B. S. & IRIZARRY, R. A. 2011. Comparing genotyping algorithms for Illumina's Infinium whole-genome SNP BeadChips. *BMC Bioinformatics,* 12**,** 68.

ROCNIK, E. F., LIU, P., SATO, K., WALSH, K. & VAZIRI, C. 2006. The novel SPARC family member SMOC-2 potentiates angiogenic growth factor activity. *J Biol Chem,* 281**,** 22855-64.

ROLLING, S. 1980. Hypodontia of permanent teeth in Danish schoolchildren. *Scand J Dent Res,* 88**,** 365-9.

ROLLING, S. & POULSEN, S. 2001. Oligodontia in Danish schoolchildren. *Acta Odontol Scand,* 59**,** 111-2.

RUFINI, A., BARLATTANI, A., DOCIMO, R., VELLETRI, T., NIKLISON-CHIROU, M. V., AGOSTINI, M. & MELINO, G. 2011. p63 in tooth development. *Biochem Pharmacol,* 82**,** 1256-61.

RUFINI, A., WEIL, M., MCKEON, F., BARLATTANI, A., MELINO, G. & CANDI, E. 2006. p63 protein is essential for the embryonic development of vibrissae and teeth. *Biochem Biophys Res Commun,* 340**,** 737-41.

RUNE, B. & SARNAS, K. V. 1974. Tooth size and tooth formation in children with advanced hypodontia. *Angle Orthod,* 44**,** 316-21.

SALAMA, F. S. & ABDEL-MEGID, F. Y. 1994. Hypodontia of primary and permanent teeth in a sample of Saudi children. *Egypt Dent J,* 40**,** 625-32.

SALINAS, C. F., JORGENSON, R. J., WRIGHT, J. T., DIGIOVANNA, J. J. & FETE, M. D. 2009. 2008 International Conference on Ectodermal Dysplasias Classification: conference report. *Am J Med Genet A,* 149A**,** 1958-69.

SARKAR, L., COBOURNE, M., NAYLOR, S., SMALLEY, M., DALE, T. & SHARPE, P. T. 2000. Wnt/Shh interactions regulate ectodermal boundary formation during mammalian tooth development. *Proc Natl Acad Sci U S A,* 97**,** 4520-4.

SARNAS, K. V. & RUNE, B. 1983. The facial profile in advanced hypodontia: a mixed longitudinal study of 141 children. *Eur J Orthod,* 5**,** 133-43.

SATOKATA, I., MA, L., OHSHIMA, H., BEI, M., WOO, I., NISHIZAWA, K., MAEDA, T., TAKANO, Y., UCHIYAMA, M., HEANEY, S., PETERS, H., TANG, Z., MAXSON, R. & MAAS, R. 2000. Msx2 deficiency in mice causes pleiotropic defects in bone growth and ectodermal organ formation. *Nat Genet,* 24**,** 391-5.

SATOKATA, I. & MAAS, R. 1994. Msx1 deficient mice exhibit cleft palate and abnormalities of craniofacial and tooth development. *Nat Genet,* 6**,** 348-56.

SCHALK-VAN DER WEIDE, Y., STEEN, W. H. & BOSMAN, F. 1992. Distribution of missing teeth and tooth morphology in patients with oligodontia. *ASDC J Dent Child,* 59**,** 133-40.

SCHALK-VAN DER WEIDE, Y., STEEN, W. H. & BOSMAN, F. 1993. Taurodontism and length of teeth in patients with oligodontia. *Journal of Oral Rehabilitation,* 20**,** 401-12.

SCHALKVANDERWEIDE, Y. & BOSMAN, F. 1996. Tooth size in relatives of individuals with oligodontia. *Archives of Oral Biology,* 41**,** 469-472.

SCHALKVANDERWEIDE, Y., STEEN, W. H. A., BEEMER, F. A. & BOSMAN, F. 1994. Reductions in Size and Left-Right Asymmetry of Teeth in Human Oligodontia. *Archives of Oral Biology,* 39**,** 935-939.

SEOW, W. K. 1993. Taurodontism of the mandibular first permanent molar distinguishes between the tricho-dento-osseous (TDO) syndrome and amelogenesis imperfecta. *Clin Genet,* 43**,** 240-6.

SHAH, R. M., BOYD, M. A. & VAKIL, T. F. 1978. Studies of permanent tooth anomalies in 7,886 Canadian individuals. II: congenitally missing, supernumerary and peg teeth. *Dent J,* 44**,** 265-8, 276.

SHAPIRA, J., CHAUSHU, S. & BECKER, A. 2000a. Prevalence of tooth transposition, third molar agenesis, and maxillary canine impaction in individuals with Down syndrome. *Angle Orthod,* 70**,** 290-6.

SHAPIRA, Y., LUBIT, E. & KUFTINEC, M. M. 1999. Congenitally missing second premolars in cleft lip and cleft palate children. *Am J Orthod Dentofacial Orthop,* 115**,** 396-400.

SHAPIRA, Y., LUBIT, E. & KUFTINEC, M. M. 2000b. Hypodontia in children with various types of clefts. *Angle Orthod,* 70**,** 16-21.

SHARPE, P. T. 2001. Neural crest and tooth morphogenesis. *Adv Dent Res,* 15**,** 4-7.

SIMONSON, M. A., WILLS, A. G., KELLER, M. C. & MCQUEEN, M. B. 2011. Recent methods for polygenic analysis of genome-wide data implicate an important effect of common variants on cardiovascular disease risk. *BMC Med Genet,* 12**,** 146.

SINGLETON, A. B. 2011. Exome sequencing: a transformative technology. *Lancet Neurol,* 10**,** 942-6.

SLAYTON, R. L., WILLIAMS, L., MURRAY, J. C., WHEELER, J. J., LIDRAL, A. C. & NISHIMURA, C. J. 2003. Genetic association studies of cleft lip and/or palate with hypodontia outside the cleft region. *Cleft Palate Craniofac J,* 40**,** 274-9.

SMAHI, A., COURTOIS, G., VABRES, P., YAMAOKA, S., HEUERTZ, S., MUNNICH, A., ISRAEL, A., HEISS, N. S., KLAUCK, S. M., KIOSCHIS, P., WIEMANN, S., POUSTKA, A., ESPOSITO, T., BARDARO, T., GIANFRANCESCO, F., CICCODICOLA, A., D'URSO, M., WOFFENDIN, H., JAKINS, T., DONNAI, D., STEWART, H., KENWRICK, S. J., ARADHYA, S., YAMAGATA, T., LEVY, M., LEWIS, R. A. & NELSON, D. L. 2000. Genomic rearrangement in NEMO impairs NF-kappaB activation and is a cause of incontinentia pigmenti. The International Incontinentia Pigmenti (IP) Consortium. *Nature,* 405**,** 466-72.

SMITH, F. J., CORDEN, L. D., RUGG, E. L., RATNAVEL, R., LEIGH, I. M., MOSS, C., TIDMAN, M. J., HOHL, D., HUBER, M., KUNKELER, L., MUNRO, C. S., LANE, E. B. & MCLEAN, W. H. 1997. Missense mutations in keratin 17 cause either pachyonychia congenita type 2 or a phenotype resembling steatocystoma multiplex. *J Invest Dermatol,* 108**,** 220-3.

SMITH, R., ZAITOUN, H., COXON, T., KARMO, M., KAUR, G., TOWNSEND, G., HARRIS, E. F. & BROOK, A. 2009. Defining new dental phenotypes using 3-D image analysis to enhance discrimination and insights into biological processes. *Arch Oral Biol,* 54 Suppl 1**,** S118-25.

SOFAER, J. A., CHUNG, C. S., NISWANDER, J. D. & RUNCK, D. W. 1971. Developmental interaction, size and agenesis among permanent maxillary incisors. *Hum Biol,* 43**,** 36-45.

STAHL, E. A., WEGMANN, D., TRYNKA, G., GUTIERREZ-ACHURY, J., DO, R., VOIGHT, B. F., KRAFT, P., CHEN, R., KALLBERG, H. J., KURREEMAN, F. A., KATHIRESAN, S., WIJMENGA, C., GREGERSEN, P. K., ALFREDSSON, L., SIMINOVITCH, K. A., WORTHINGTON, J., DE BAKKER, P. I., RAYCHAUDHURI, S. & PLENGE, R. M. 2012. Bayesian inference analyses of the polygenic architecture of rheumatoid arthritis. *Nat Genet,* 44**,** 483-9.

STANKIEWICZ, P. & LUPSKI, J. R. 2010. Structural variation in the human genome and its role in disease. *Annu Rev Med,* 61**,** 437-55.

STENVIK, A., ZACHRISSON, B. U. & SVATUN, B. 1972. Taurodontism and concomitant hypodontia in siblings. *Oral Surg Oral Med Oral Pathol,* 33**,** 841-5.

STOCKTON, D. W., DAS, P., GOLDENBERG, M., D'SOUZA, R. N. & PATEL, P. I. 2000. Mutation of PAX9 is associated with oligodontia. *Nat Genet,* 24**,** 18-9.

STRACHAN, T. & READ, A. P. 2010. *Human molecular genetics,* New York, Garland Science.

STRITZEL, F., SYMONS, A. L. & GAGE, J. P. 1990. Agenesis of the second premolar in males and females: distribution, number and sites affected. *J Clin Pediatr Dent,* 15**,** 39-41.

SUAREZ, B. K. & SPENCE, M. A. 1974. The genetics of hypodontia. *J Dent Res,* 53**,** 781-5.

SUZUKI, K., HU, D., BUSTOS, T., ZLOTOGORA, J., RICHIERI-COSTA, A., HELMS, J. A. & SPRITZ, R. A. 2000. Mutations of PVRL1, encoding a cell-cell adhesion molecule/herpesvirus receptor, in cleft lip/palate-ectodermal dysplasia. *Nat Genet,* 25**,** 427-30.

SUZUKI, Y., JEZEWSKI, P. A., MACHIDA, J., WATANABE, Y., SHI, M., COOPER, M. E., VIET LE, T., NGUYEN, T. D., HAI, H., NATSUME, N., SHIMOZATO, K., MARAZITA, M. L. & MURRAY, J. C. 2004. In a Vietnamese population, MSX1 variants contribute to cleft lip and palate. *Genet Med,* 6**,** 117-25.

SVINHUFVUD, E., MYLLARNIEMI, S. & NORIO, R. 1988. Dominant inheritance of tooth malpositions and their association to hypodontia. *Clin Genet,* 34**,** 373-81.

SWINNEN, S., BAILLEUL-FORESTIER, I., ARTE, S., NIEMINEN, P., DEVRIENDT, K. & CARELS, C. 2008. Investigating the etiology of multiple tooth agenesis in three sisters with severe oligodontia. *Orthod Craniofac Res,* 11**,** 24-31.

TALLON-WALTON, V., MANZANARES-CESPEDES, M. C., ARTE, S., CARVALHO-LOBATO, P., VALDIVIA-GANDUR, I., GARCIA-SUSPERREGUI, A., VENTURA, F. & NIEMINEN, P. 2007. Identification of a novel mutation in the PAX9 gene in a family affected by oligodontia and other dental anomalies. *Eur J Oral Sci,* 115**,** 427-32.

TAO, R., JIN, B., GUO, S. Z., QING, W., FENG, G. Y., BROOKS, D. G., LIU, L., XU, J., LI, T., YAN, Y. & HE, L. 2006. A novel missense mutation of the EDA gene in a Mongolian family with congenital hypodontia. *J Hum Genet,* 51**,** 498-502.

THEIS, J. L., SHARPE, K. M., MATSUMOTO, M. E., CHAI, H. S., NAIR, A. A., THEIS, J. D., DE ANDRADE, M., WIEBEN, E. D., MICHELS, V. V. & OLSON, T. M. 2011. Homozygosity mapping and exome sequencing reveal GATAD1 mutation in autosomal recessive dilated cardiomyopathy. *Circ Cardiovasc Genet,* 4**,** 585-94.

THESLEFF, I. 1995. Tooth morphogenesis. *Adv Dent Res,* 9**,** 12.

THESLEFF, I. 2000. Genetic basis of tooth development and dental defects. *Acta Odontologica Scandinavica,* 58**,** 191-194.

THESLEFF, I. 2003. Developmental biology and building a tooth. *Quintessence Int,* 34**,** 613-20.

THESLEFF, I. 2006. The genetic basis of tooth development and dental defects. *Am J Med Genet A,* 140**,** 2530-5.

THESLEFF, I. & ABERG, T. 1997. Tooth morphogenesis and the differentiation of ameloblasts. *Ciba Found Symp,* 205**,** 3-12; discussion 12-7.

THESLEFF, I. & SHARPE, P. 1997. Signalling networks regulating dental development. *Mech Dev,* 67**,** 111-23.

THOMAS, B. L., TUCKER, A. S., QUI, M., FERGUSON, C. A., HARDCASTLE, Z., RUBENSTEIN, J. L. & SHARPE, P. T. 1997. Role of Dlx-1 and Dlx-2 genes in patterning of the murine dentition. *Development,* 124**,** 4811-8.

THOMPSON, G. W. & POPOVICH, F. 1974. Probability of congenitally missing teeth: results in 1,191 children in the Burlington Growth centre in Toronto. *Community Dent Oral Epidemiol,* 2**,** 26-32.

THYSSEN, J. P., GODOY-GIJON, E. & ELIAS, P. M. 2013. Ichthyosis vulgaris: the filaggrin mutation disease. *Br J Dermatol,* 168**,** 1155-66.

TOWNSEND, G., HARRIS, E. F., LESOT, H., CLAUSS, F. & BROOK, A. 2009. Morphogenetic fields within the human dentition: a new, clinically relevant synthesis of an old concept. *Arch Oral Biol,* 54 Suppl 1**,** S34-44.

TOWNSEND, G., ROGERS, J., RICHARDS, L. & BROWN, T. 1995. Agenesis of permanent maxillary lateral incisors in South Australian twins. *Aust Dent J,* 40**,** 186-92.

TOWNSEND, G. C., RICHARDS, L., HUGHES, T., PINKERTON, S. & SCHWERDT, W. 2005. Epigenetic influences may explain dental differences in monozygotic twin pairs. *Aust Dent J,* 50**,** 95-100.

TUCKER, A. & SHARPE, P. 2004. The cutting-edge of mammalian development; how the embryo makes teeth. *Nat Rev Genet,* 5**,** 499-508.

TUCKER, A. S., MATTHEWS, K. L. & SHARPE, P. T. 1998. Transformation of tooth type induced by inhibition of BMP signaling. *Science,* 282**,** 1136-8.

TUCKER, A. S., YAMADA, G., GRIGORIOU, M., PACHNIS, V. & SHARPE, P. T. 1999. Fgf-8 determines rostral-caudal polarity in the first branchial arch. *Development,* 126**,** 51-61.

UNER, O., YÜCEL-EROĞLU, E. & KARACA, I. 1994. Delayed calcification and congenitally missing teeth. Case report. *Aust Dent J,* 39**,** 168-71.

VAINIO, S., KARAVANOVA, I., JOWETT, A. & THESLEFF, I. 1993. Identification of BMP-4 as a signal mediating secondary induction between epithelial and mesenchymal tissues during early tooth development. *Cell,* 75**,** 45-58.

VAN DEN BOOGAARD, M. J., CRETON, M., BRONKHORST, Y., VAN DER HOUT, A., HENNEKAM, E., LINDHOUT, D., CUNE, M. & PLOOS VAN AMSTEL, H. K. 2012. Mutations in WNT10A are present in more than half of isolated hypodontia cases. *J Med Genet,* 49**,** 327-31.

VAN DEN BOOGAARD, M. J., DORLAND, M., BEEMER, F. A. & VAN AMSTEL, H. K. 2000. MSX1 mutation is associated with orofacial clefting and tooth agenesis in humans. *Nat Genet,* 24**,** 342-3.

VANNAHME, C., GOSLING, S., PAULSSON, M., MAURER, P. & HARTMANN, U. 2003. Characterization of SMOC-2, a modular extracellular calcium-binding protein. *Biochem J,* 373**,** 805-14.

VASTARDIS, H. 2000. The genetics of human tooth agenesis: new discoveries for understanding dental anomalies. *Am J Orthod Dentofacial Orthop,* 117**,** 650-6.

VASTARDIS, H., KARIMBUX, N., GUTHUA, S. W., SEIDMAN, J. G. & SEIDMAN, C. E. 1996. A human MSX1 homeodomain missense mutation causes selective tooth agenesis. *Nat Genet,* 13**,** 417-21.

WALLIS, D. & MUENKE, M. 2000. Mutations in holoprosencephaly. *Hum Mutat,* 16**,** 99-108.

WANG, H., ZHAO, S., ZHAO, W., FENG, G., JIANG, S., LIU, W., LI, S., XUE, H. & HE, L. 2000. Congenital absence of permanent teeth in a six-generation Chinese kindred. *Am J Med Genet,* 90**,** 193-8.

WANG, K., LI, M. & HAKONARSON, H. 2010. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res,* 38**,** e164.

WANG, Y., KONG, H., MUES, G. & D'SOUZA, R. 2011. Msx1 mutations: how do they cause tooth agenesis? *J Dent Res,* 90**,** 311-6.

WANG, Y., WU, H., WU, J., ZHAO, H., ZHANG, X., MUES, G., D'SOUZA, R. N., FENG, H. & KAPADIA, H. 2009. Identification and functional analysis of two novel PAX9 mutations. *Cells Tissues Organs,* 189**,** 80-7.

WEI, J., XUE, Y., WU, L., MA, J., YI, X., ZHANG, J., LU, B., LI, C., SHI, D., SHI, S., FENG, X. & CAI, T. 2012. Analysis of large phenotypic variability of EEC and SHFM4 syndromes caused by K193E mutation of the TP63 gene. *PLoS One,* 7**,** e35337.

WHITTINGTON, B. R. & DURWARD, C. S. 1996. Survey of anomalies in primary teeth and their correlation with the permanent dentition. *N Z Dent J,* 92**,** 4-8.

WINNIER, G., BLESSING, M., LABOSKY, P. A. & HOGAN, B. L. 1995. Bone morphogenetic protein-4 is required for mesoderm formation and patterning in the mouse. *Genes Dev,* 9**,** 2105-16.

WONG, F. K. & GUSTAFSSON, B. 2000. Popliteal pterygium syndrome in a Swedish family--clinical findings and genetic analysis with the van der Woude syndrome locus at 1q32-q41. *Acta Odontol Scand,* 58**,** 85-8.

WONG, F. K., HAGBERG, C., KARSTEN, A., LARSON, O., GUSTAVSSON, M., HUGGARE, J., LARSSON, C., TEH, B. T. & LINDER-ARONSON, S. 2000. Linkage analysis of candidate regions in Swedish nonsyndromic cleft lip with or without cleft palate families. *Cleft Palate Craniofac J,* 37**,** 357-62.

WOOLF, C. M. 1971. Missing maxillary lateral incisors: a genetic study. *Am J Hum Genet,* 23**,** 289-96.

WORLD HEALTH ORGANISATION. 1962. Standardisation of reporting of dental disease and conditions. 6. The assessment of handicapping dentofacial anomalies. *Technical report series*, No.242, Geneva, WHO, 1962.

WU, Q. H., SHI, H. R., LIU, B. C., ZHAO, Z. H., JIANG, M., LU, N. & KONG, X. D. 2012. [Detection of ED1 gene mutations in six pedigrees with hypohidrotic ectodermal dysplasia]. *Zhonghua Yi Xue Yi Chuan Xue Za Zhi,* 29**,** 447-51.

XUAN, K., JIN, F., LIU, Y. L., YUAN, L. T., WEN, L. Y., YANG, F. S., WANG, X. J., WANG, G. H. & JIN, Y. 2008. Identification of a novel missense mutation of MSX1 gene in Chinese family with autosomal-dominant oligodontia. *Arch Oral Biol,* 53**,** 773-9.

YAMADA, G., MANSOURI, A., TORRES, M., STUART, E. T., BLUM, M., SCHULTZ, M., DE ROBERTIS, E. M. & GRUSS, P. 1995. Targeted mutation of the murine goosecoid gene results in craniofacial defects and neonatal death. *Development,* 121**,** 2917-22.

YANG, A., SCHWEITZER, R., SUN, D., KAGHAD, M., WALKER, N., BRONSON, R. T., TABIN, C., SHARPE, A., CAPUT, D., CRUM, C. & MCKEON, F. 1999. p63 is essential for regenerative proliferation in limb, craniofacial and epithelial development. *Nature,* 398**,** 714-8.

YASUDA, I., YASUDA, M., SUMIDA, H., TSUSAKI, H., ARIMA, A., IHARA, T., KUBOTA, S., ASAOKA, K., TSUGA, K. & AKAGAWA, Y. 2005. In utero and lactational exposure to 2,3,7,8-tetrachlorodibenzo-p-dioxin (TCDD) affects tooth development in rhesus monkeys. *Reprod Toxicol,* 20**,** 21-30.

YONEZU, T., HAYASHI, Y., SASAKI, J. & MACHIDA, Y. 1997. Prevalence of congenital dental anomalies of the deciduous dentition in Japanese children. *Bull Tokyo Dent Coll,* 38**,** 27-32.

YOUNG, I. D. 2005. *Medical genetics,* Oxford, Oxford University Press.

ZHANG, Y. D., CHEN, Z., SONG, Y. Q., LIU, C. & CHEN, Y. P. 2005. Making a tooth: growth factors, transcription factors, and stem cells. *Cell Res,* 15**,** 301-16.

ZHANG, H., QUAN, C., SUN, L. D., LV, H. L., GAO, M., ZHOU, F. S., XIAO, F. L., FANG, Q. Y., SHEN, Y. J., ZHOU, L., YANG, S. & ZHANG, X. J. 2009. A novel frameshift mutation of the EDA1 gene in a Chinese Han family with X-linked hypohidrotic ectodermal dysplasia. *Clin Exp Dermatol,* 34**,** 74-6.

ZHAO, J., HU, Q., CHEN, Y., LUO, S., BAO, L. & XU, Y. 2007. A novel missense mutation in the paired domain of human PAX9 causes oligodontia. *Am J Med Genet A,* 143A**,** 2592-7.

ZILBERMAN, Y., COHEN, B. & BECKER, A. 1990. Familial trends in palatal canines, anomalous lateral incisors, and related phenomena. *Eur J Orthod,* 12**,** 135-9.

ZOUVELOU, V., LUDER, H. U., MITSIADIS, T. A. & GRAF, D. 2009. Deletion of BMP7 affects the development of bones, teeth, and other ectodermal appendages of the orofacial complex. *J Exp Zool B Mol Dev Evol,* 312B**,** 361-74.

Web sites

ClustalW, http://www.ebi.ac.uk/tools/clustalw2/index.html. (accessed date: 21/07/2013).

Ensembl, http://www.ensembl.org/index.html. (accessed date: 21/07/2013).

Gene expression in tooth, http://www.bite-it.helsinki.fi. (accessed date: 21/07/2013).

Multalin Interface Page (online), http://bioinfo.genotoul.fr/multalin/multalin.html. (accessed date: 21/07/2013).

National Center for Biotechnology Information, http://www.ncbi.nlm.nih.gov/. (accessed date: 21/07/2013).

Prime3, http://www.frodo.wi.mit.edu/primer3/input.html. (accessed date: 21/07/2013).

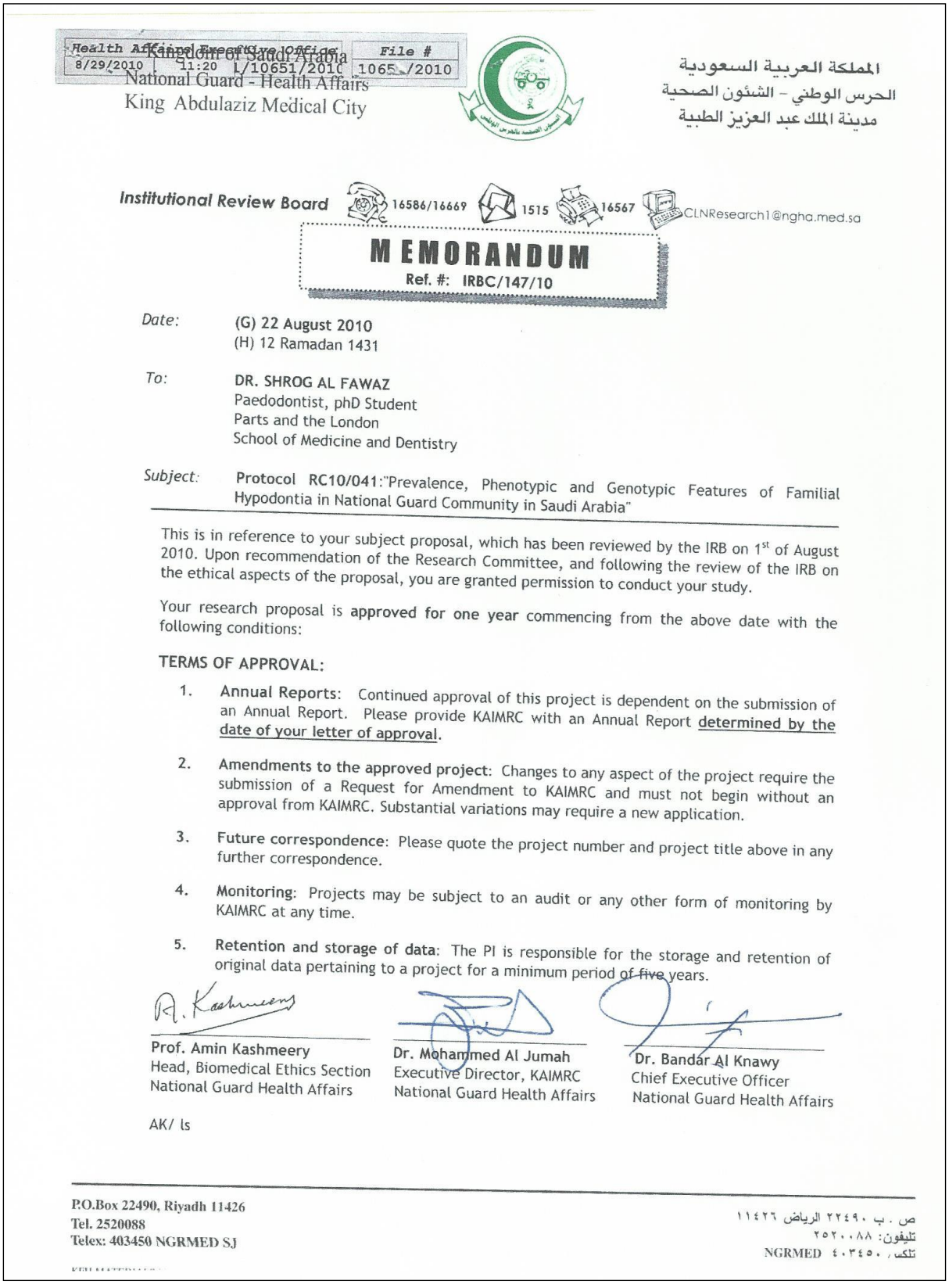
The Marshfield Clinic Research Foundation (MCRF) http://www.marshfieldclinic.org/research/pages/index.aspx. (accessed date: 21/07/2013).

The University of California Santa Cruz UCSC Genome Browser http://www.genome.ucsc.edu/. (accessed date: 21/07/2013).

The 1000 Genome Browser, http://www.browser.1000genomes.org/index.html. (accessed date: 21/07/2013).

Appendix A:

Figure A1: A copy of the project approval memo (from the ethics committee of the National Guard, Health Affairs).



Health Affairs Executive Office
8/29/2010 | 11:20 | I/10651/2010
File # 1065 /2010

Kingdom of Saudi Arabia
National Guard - Health Affairs
King Abdulaziz Medical City

المملكة العربية السعودية
الحرس الوطني - الشئون الصحية
مدينة الملك عبد العزيز الطبية

Institutional Review Board 16586/16669 1515 16567 CLNResearch1@ngha.med.sa

MEMORANDUM
Ref. #: IRBC/147/10

Date: (G) 22 August 2010
(H) 12 Ramadan 1431

To: DR. SHROG AL FAWAZ
Paedodontist, phD Student
Parts and the London
School of Medicine and Dentistry

Subject: Protocol RC10/041:"Prevalence, Phenotypic and Genotypic Features of Familial Hypodontia in National Guard Community in Saudi Arabia"

This is in reference to your subject proposal, which has been reviewed by the IRB on 1st of August 2010. Upon recommendation of the Research Committee, and following the review of the IRB on the ethical aspects of the proposal, you are granted permission to conduct your study.

Your research proposal is **approved for one year** commencing from the above date with the following conditions:

TERMS OF APPROVAL:

1. **Annual Reports:** Continued approval of this project is dependent on the submission of an Annual Report. Please provide KAIMRC with an Annual Report **determined by the date of your letter of approval**.
2. **Amendments to the approved project:** Changes to any aspect of the project require the submission of a Request for Amendment to KAIMRC and must not begin without an approval from KAIMRC. Substantial variations may require a new application.
3. **Future correspondence:** Please quote the project number and project title above in any further correspondence.
4. **Monitoring:** Projects may be subject to an audit or any other form of monitoring by KAIMRC at any time.
5. **Retention and storage of data:** The PI is responsible for the storage and retention of original data pertaining to a project for a minimum period of five years.

Prof. Amin Kashmeery
Head, Biomedical Ethics Section
National Guard Health Affairs

Dr. Mohammed Al Jumah
Executive Director, KAIMRC
National Guard Health Affairs

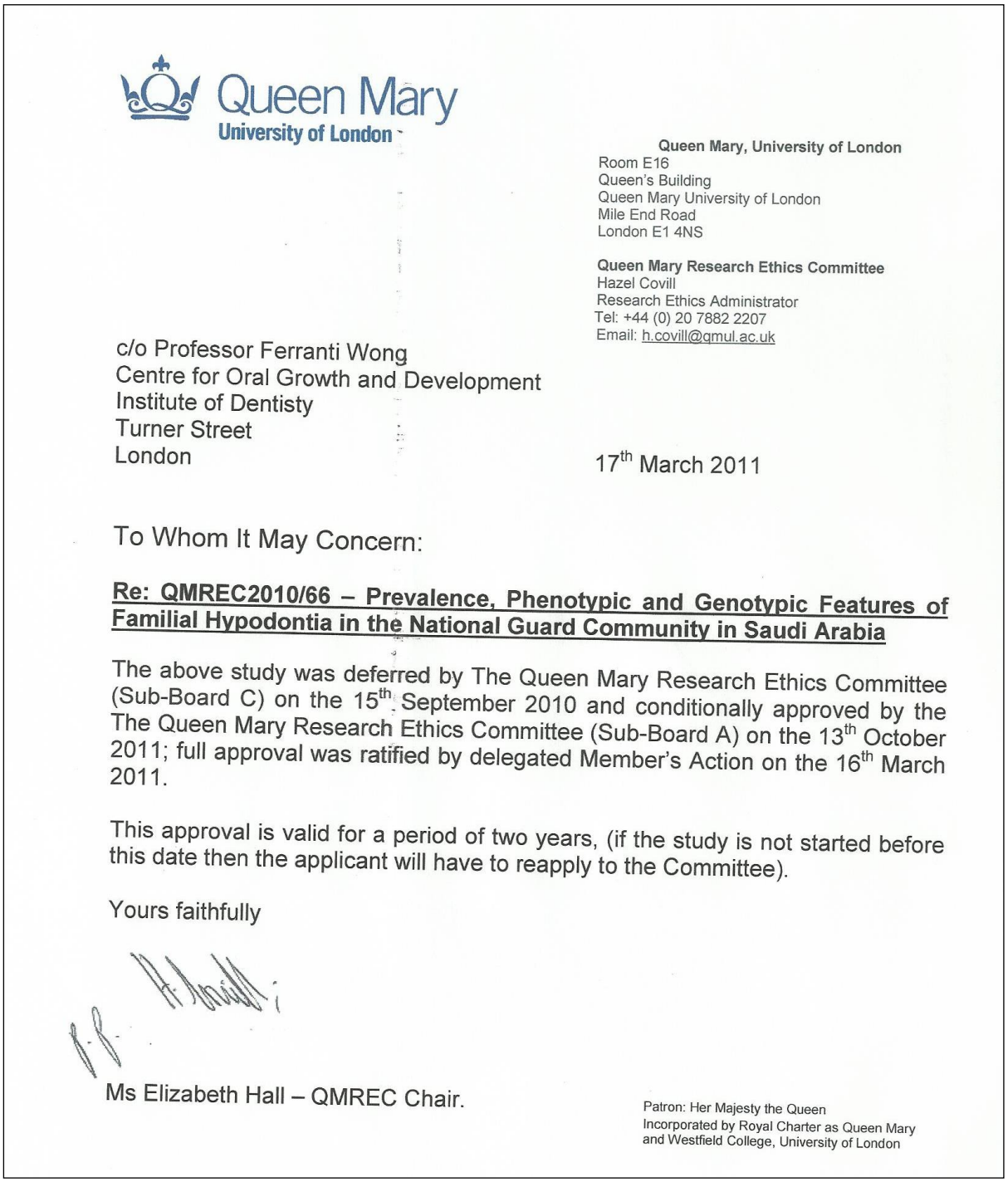
Dr. Bandar Al Knawy
Chief Executive Officer
National Guard Health Affairs

AK/ ls

P.O.Box 22490, Riyadh 11426
Tel. 2520088
Telex: 403450 NGRMED SJ

ص . ب ٢٢٤٩٠ الرياض ١١٤٢٦
تليفون: ٢٥٢٠٠٨٨
تلكس / ٤٠٣٤٥٠ NGRMED

Figure A2: A copy of the project approval letter (from the Queen Mary University of London (QMUL) research ethics committee)



Queen Mary
University of London

Queen Mary, University of London
Room E16
Queen's Building
Queen Mary University of London
Mile End Road
London E1 4NS

Queen Mary Research Ethics Committee
Hazel Covill
Research Ethics Administrator
Tel: +44 (0) 20 7882 2207
Email: h.covill@qmul.ac.uk

c/o Professor Ferranti Wong
Centre for Oral Growth and Development
Institute of Dentisty
Turner Street
London

17th March 2011

To Whom It May Concern:

Re: QMREC2010/66 – Prevalence, Phenotypic and Genotypic Features of Familial Hypodontia in the National Guard Community in Saudi Arabia

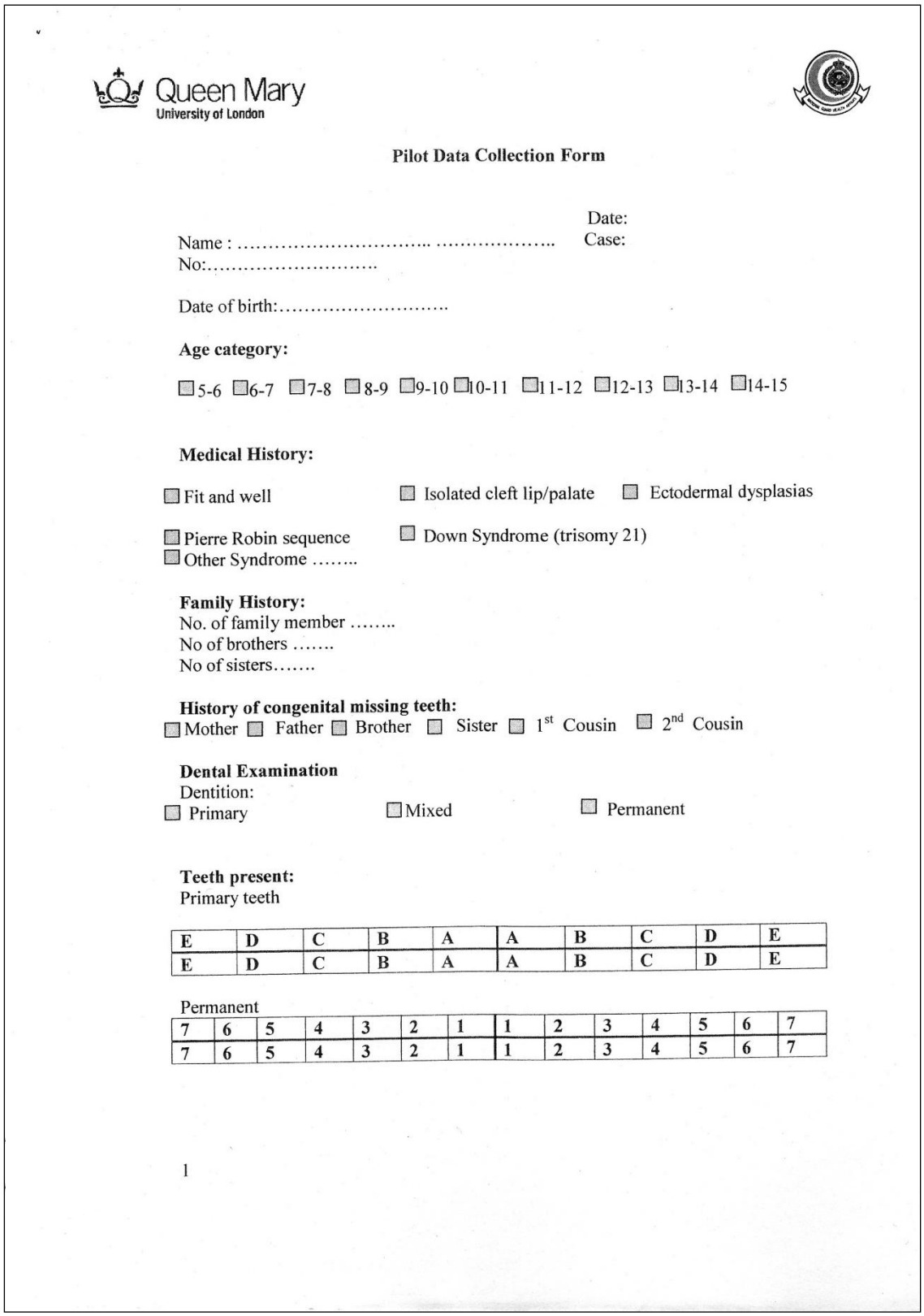
The above study was deferred by The Queen Mary Research Ethics Committee (Sub-Board C) on the 15th September 2010 and conditionally approved by the The Queen Mary Research Ethics Committee (Sub-Board A) on the 13th October 2011; full approval was ratified by delegated Member's Action on the 16th March 2011.

This approval is valid for a period of two years, (if the study is not started before this date then the applicant will have to reapply to the Committee).

Yours faithfully

Ms Elizabeth Hall – QMREC Chair.

Figure A3: The clinical questionnaire



Queen Mary
University of London

Pilot Data Collection Form

Date:
Name : …….. …….. Case:
No:……

Date of birth:……

Age category:

☐5-6 ☐6-7 ☐7-8 ☐8-9 ☐9-10 ☐10-11 ☐11-12 ☐12-13 ☐13-14 ☐14-15

Medical History:

☐ Fit and well ☐ Isolated cleft lip/palate ☐ Ectodermal dysplasias

☐ Pierre Robin sequence ☐ Down Syndrome (trisomy 21)
☐ Other Syndrome ……..

Family History:
No. of family member ……..
No of brothers …….
No of sisters…….

History of congenital missing teeth:
☐ Mother ☐ Father ☐ Brother ☐ Sister ☐ 1st Cousin ☐ 2nd Cousin

Dental Examination
Dentition:
☐ Primary ☐ Mixed ☐ Permanent

Teeth present:
Primary teeth

E	D	C	B	A	A	B	C	D	E
E	D	C	B	A	A	B	C	D	E

Permanent

7	6	5	4	3	2	1	1	2	3	4	5	6	7
7	6	5	4	3	2	1	1	2	3	4	5	6	7

1

History of Extraction:

Primary teeth

E	D	C	B	A	A	B	C	D	E
E	D	C	B	A	A	B	C	D	E

Permanent

7	6	5	4	3	2	1	1	2	3	4	5	6	7
7	6	5	4	3	2	1	1	2	3	4	5	6	7

Missing teeth:

☐ None

☐ Primary teeth

E	D	C	B	A	A	B	C	D	E
E	D	C	B	A	A	B	C	D	E

☐ Permanent

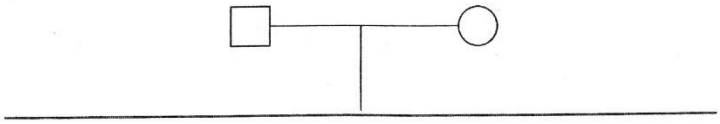
7	6	5	4	3	2	1	1	2	3	4	5	6	7
7	6	5	4	3	2	1	1	2	3	4	5	6	7

Dental Anomalies

☐ None
☐ Delayed eruption of teeth
☐ Infraocluded primary molar
☐ Microdontia
  ☐ Peg-shaped permanent upper lateral incisors.
  ☐ Generalised
☐ Malposition of teeth:
  ☐ Ectopic maxillary canines
  ☐ Ectopic eruption of other teeth
☐ Enamel hypoplasia, hypocalcification
☐ Taurodontism
☐ Short roots of teeth
☐ Other

2

Family Pedigree



3

Appendix B: DNA extraction

Protocol used: Purifying of DNA from Whole Blood Using the QIAamp Blood Midi Kit

Procedure

- 100µl QIAGEN Protease was loaded with a pipet in to the bottom of a 15 ml centrifuge tube.
- 1 ml of blood sample was added and mixed briefly.
- 1.2 ml Buffer AL was added and mixed thoroughly by inverting and shaking the tube.
- The sample was incubated at 70°C for 10 min.
- 1 ml ethanol (100%) was added to the sample and mixed by inverting and shaking the tube.
- All the solution was carefully transferred from the tube sample on to the QIAampMidi column placed in a 15 ml centrifuge tube. And centrifuged at 1850xg (3000 rpm) for 3 min.
- The filtrate was discarded, and then 2 ml Buffer AW1 was added to the QIAampMidi column and centrifuged at 4500 x g (5000rpm) for 15 min.
- The QIAampMidi column was placed in a clean 15 ml centrifuge tube. And 200 µl distilled water (equilibrated to room temperature was pipetted directly on to the membrane of the QIAampMidi column and incubated at room temperature for 5 min, then centrifuged at 4500 x g (5000 rpm) for 2 min.
- Finally for maximum yield , 200 µl distilled water (equilibrated to room temperature was pipetted directly on to the membrane of the QIAampMidi column and incubated at room temperature for 5 min, then centrifuged at 4500 x g (5000 rpm) for 2 min.

Protocol used: Manual purification of DNA from 0.5 ml of Oragene. DNA/saliva

Procedure

- Oragene DNA/saliva sample was mixed gently and incubated at 50° C in a water incubator for a minimum of 1 hour.
- 500 µl of the mixed Oragene DNA/saliva sample was transferred to a 1.5 ml micro-centrifuge tube first, then 20 µl of Purifier was added to it and mixed by vortexing for a few seconds.
- The sample was incubated on ice for 10 min, then centrifuged at room temperature for 5 min at 13,000 rpm (15, 000 x g).
- The clear supernatant was carefully transferred in to a fresh micro-centrifuge tube and then 500 µl of room temperature 100% ethanol was added and gently mixed.
- The sample then was left at room temperature for 10 to allow the DNA to fully precipitate. Then centrifuged at room temperature for 2 min at 13,000 rpm (15.000 x g).
- The supernatant was carefully removed and then 250 µl of 70% ethanol was added to the DNA pellet, left for 1 min then completely removed without disturbing the pellet.
- Finally 100 µl of distilled water was added to the pellet and the sample then vortexed for DNA hydration.

Appendix C: Primers sequences

Table C1: Primers used for amplification and sequencing of genomic DNA (Sanger sequencing).

Gene	Forward Primer Sequence 5'-3'	Reverse Primer Sequence 5'-3'	Product size (bP)	Reference
MSX1	AGTCGCCAGAGGAAAGTTTC	CCTTGGCCCCCGGCTTCCTGTGGTC	650	This study
	CCAAGGAGAGCGCCCTGGCGCCCT	CTCCCTCTGCGCCTGGGTTCTGGCT	277	
	GCAAACACAAGACGAACCGTAAGCC	CTTGTAGTCTCTTTGCCTTGGCGCG	200	
	CAATGCTTCTCTCTTAACCCTTGCTT	CTGGCCCCACAGGTGGGCGGACCTG	520	
	ACTTGGCGGCACTCAATATC	CAGGGAGCAAAGAGGTGAAA	698	
PAX9	AGGCAGCTGTCCCAAGCAGCG	GGAGGGCACATTGTACTTGTCGC	410	This study
	ATCCGACCGTGTGACATCAGCC	GAGCCCCTACCTTGGTCGGTG	525	
	GGGAGTAAAACTTCACCAGGC	CCACCTGGCCTGACCCTC	370	
	GGAGAGTAGAGTCAGAGCATTGCTG	GAGACCTGGGAATTGGGGGA	590	
KREMEN1	TCCATGCTTTCGTTTTGG	AGCGAGACCCCATCTCTACA	208	This study
	TGAGTAGCTGGGACCACTGA	CTCAGAGCGCTTGCACTAAA	193	
	GCTGGGATCACAGGAGTTGT	TCCCTTGCTTAGCTCATCGT	540	
	TTTTTAACCATTACGATGAGCTAAG	TCCAAGGGTAGGATCAGCAG	217	
	CGTGGGTCACGACCTAAGTG	GCTGCTCCTGCCACTCAC	540	
	CTCCCGCTTCAGGACCTTC	AGGGTCTTGGAAGACTCAAGG	520	

KREMEN1	TTAAATCCAAGTGTCAGTGCAG	AAGCCCATGAAGTGAACCAG	500	This study
	GATGGGGCCACTCAGTACTTT	ATCCTGGCTAACACGGTTGA	350	
	TTGGCATGATGTGTGCAGT	CAACTGGTTCATTGCTTTTCA	232	
	CCTGCCAGAGGTCAGGTAGTT	AATTCTCATGGGGCTTACCC	561	
	TGCTGATTTCCTTAGCACCTC	TACTTCTGGGGACCCATTCA	256	
	AATTTGTCCCAGCAGCAATAA	GAGTTACCAGAGGGCAGGAA	532	
	GAGGAAACCCACTTGTGGAA	GGATTTAGTATTGAAGAAAACTAGGAA	212	
	AACATGGTTTTTAGAGGACA	GCAACATCTGGCTGTGAAAC	278	
	GTGATGACTCTGTGGCTGTG	GTACCAATTCCTCCTCCCGCTTC	218	
	TGCATTTTGCTGAGATAACTTAAAA	CAGTCACTGCCATCTAATTCCA	238	
	CGGCTGTGGTTACACAAATCT	GTTTGTGTTTCTACTGCACTTCA	249	
	TGAGATTTTCTCCCCCTTAAAA	GCATCATCAGGGATGTCAGA	245	
STRN3	TCATGAGTCAAGTGCAACAGTTT	TGCCAAATACAGAGTACAATGG	287	This study
SMOC2	GAGGACCACATCGTTCTTGG	TGACTGTGTGTCGGTGTGTG	398	This study
ZFYVE26	GACTGGTTAATGGTGCCTGTG	CCCTGGGCAAGTTTCCTAAC	973	This study
	GGATCCTCCAAGCTCCTCTT	ATGCAAAGCAAAACCCAGAC	439	
KLK10	AGCTCTCACGTTGGCACATT	GGCAGAGAATGGGGATAGGT	400	This study
FOXE1	CTACAGCTACATCGCGCTCA	CTGGTAGCCGGTGGTGGTAG	622	This study

Table C2: Primers used for amplification and sequencing of genomic DNA (Microsatellite marker)

Marker	Location	Chromosome	Forward Primer (5'-3')	Reverse Primer (5'-3')	Allele Size Range (bp)	Het Score
D19S925	21361456 - 21361759	19	CTGAAACGATTCCAAATGATG	GTATGTTGAACCAGCCTTGC	232 - 274	0.67
D19S568	22495123 - 22495405	19	TGAGTCTGCTGAGACCAAAGTTAG	ATAATGTAGCCTTGTCCTGGAATAG	249 - 275	0.62
D19S434	22767818 - 22768141	19	GCCCACCATATTACTGTGGA	TTGACTAAAGCCACCACACA	264 - 280	0.52
D19S1036	23652786 - 23653103	19	AAGCAGCATCCAGCACAC	AGGTGACAGAGCAAGATTCTT	204 - 216	0.57
D19S419	28291180 - 28291487	19	GATTATCGGGGCAGGT	CATTAAATATGAAATTCAGGTAAGC	165 -169	0.58
AFMA247XE1	27825880-27826212	22	GGAACAGTTTGTGCCCC	CAGGCTTCATACCCAGACC	213-235	0.67
UT5900	29023773-29024204	22	ACAGCCTGGGTGACAGAGT	GGTGAAGCCCTGTGCTATG	178	0.49
UT5072	29242270-29242629	22	AGCCCAGGAGTTCTCTGTC	TCTCATCTCACACACCATGC	432	0.66
AFMA043TF9	30551121-30551268	22	CTGTGCCATCTCTAGTGGAC	GGGGTATGCGTGTACG	65-87	0.67
AFM106XD2	30576610-30576822	22	GGCTACAGCCAGACCC	GCCATAGGCAATGAAGTG	194-206	0.74
GATA190C03	31289840-31290170	22	CCATCCAAAATATATCTATCTATCC	ACCGAGAGAGAGAGAGAGATTG	203-231	0.63
AFM225XF6	31533926-31534254	22	GCTCCAGCCTATCAGGATG	GATTCCAGATCACAAAACTGGT	208-220	0.82
UT1963	31844607-31845073	22	ATATCAGCTACTCGAGCGG	AAAGAGCTGCTAACACTGCA	293	0.49
AFMB360ZF9	31998946-31999257	22	TCAAGCTGCTATGTTTGTGG	TTTGAGCAAGGCTTCCC	224-272	0.48
AFMB294YD5	32635470-32635866	22	TGTCTCCCAATCAACATAGG	TTTAGAAACAGCAGCCAGAA	155-167	0.80
AFM238WC11	32661262-32661560	22	GCTCAAACAACTGCAAGTGGTTCT	TGTCTTCAAAATCATTTGGGGTCC	135-151	0.83

Appendix D: Molecular biology materials

1x TBE

0.89 M tris base
0.89 M boric acid
0.2 M EDTA, pH 8.0

DNA loading dye

50% (v/v) glycerol
50% dH_2O
1% (w/v) orange G

Appendix E: Family I

Table E1: Rare variants in batch 1

	Full Gene Name	AA Change	Func	Exonic Func	Call_341
MYOM3	myomesin family, member 3	uc001bim.3:c.C280T:p.L94F	exonic	nonsynonymous SNV	het
SMPDL3B	sphingomyelin phosphodiesterase, acid-like 3B	uc010ofr.1:c.G727C:p.G243R	exonic; splicing	nonsynonymous SNV	het
TTC39A	tetratricopeptide repeat domain 39A	uc001cso.1:c.G1102T:p.G368C	exonic	nonsynonymous SNV	het
C8B	complement component 8, beta polypeptide	uc001cyp.2:c.C829T:p.H277Y	exonic	nonsynonymous SNV	het
OR6K6	olfactory receptor, family 6, subfamily K, member 6	uc001fsw.1:c.C929T:p.P310L	exonic	nonsynonymous SNV	het
LBR	lamin B receptor	uc001hoy.2:c.G224A:p.R75H	exonic	nonsynonymous SNV	het
OBSCN	obscurin, cytoskeletal calmodulin and titin-interacting RhoGEF	uc001hsr.1:c.G3574A:p.E1192K	exonic	nonsynonymous SNV	het
PLA2G12B	phospholipase A2, group XIIB	uc001jtf.1:c.C72G:p.S24R	exonic	nonsynonymous SNV	het
SFTPA1	surfactant protein A1	uc010qlt.1:c.C418T:p.R140C	exonic	nonsynonymous SNV	het
SH2D4B	SH2 domain containing 4B	uc001kcm.1:c.531_533del:p.177_178del	exonic	nonframeshift deletion	het
JAKMIP3	Janus kinase and microtubule interacting protein 3	uc001lkx.3:c.660_661del:p.220_221del	exonic	frameshift deletion	het
TUBGCP2	tubulin, gamma complex associated protein 2	uc009ybk.1:c.C1916T:p.T639M	exonic	nonsynonymous SNV	het
UBGCP2	tubulin, gamma complex	uc001lmf.1:c.C612A:p.S204R	exonic	nonsynonymous	Het

	associated protein 2			SNV	
FCHSD2	FCH and double SH3 domains 2	uc001oth.3:c.A140G:p.N47S	exonic	nonsynonymous SNV	het
MAML2	mastermind-like 2 (Drosophila)	uc001pfw.1:c.1804_1805insCAGCAA:p.Q602delinsQQQ	exonic	nonframeshift insertion	het
PLEKHA9	NA	uc001rom.1:c.G229A:p.A77T	exonic	nonsynonymous SNV	het
KRT2	keratin 2	uc001sat.2:c.G407A:p.G136E	exonic	nonsynonymous SNV	het
KIAA0748	KIAA0748	uc001sgl.3:c.G1007T:p.R336L	exonic	nonsynonymous SNV	het
OR6C65	olfactory receptor, family 6, subfamily C, member 65	uc010spl.1:c.82_83insACAAGTTGTGATATTCTTCT:p.F28_M29delinsYKLX	exonic	stopgain SNV	het
KIAA1002	NA	uc009zpn.1:c.G506A:p.R169H	exonic	nonsynonymous SNV	het
2	SLIT-ROBO Rho GTPase activating protein 1	uc001srv.2:c.A2947G:p.M983V	exonic	nonsynonymous SNV	het
ZDHHC17	zinc finger, DHHC-type containing 17	uc001syk.1:c.T752C:p.V251A	exonic	nonsynonymous SNV	het
C12orf51	chromosome 12 open reading frame 51	uc001tts.2:c.A397C:p.I133L	exonic	nonsynonymous SNV	het
TNFSF11	tumor necrosis factor (ligand) superfamily, member 11	uc001uyu.2:c.G937A:p.V313I	exonic	nonsynonymous SNV	het
OR4K14	olfactory receptor, family 4, subfamily K, member 14	uc010tky.1:c.G449C:p.W150S	exonic	nonsynonymous SNV	het
STRN3	striatin, calmodulin binding protein 3	uc001wqu.2:c.G287A:p.R96Q	exonic	nonsynonymous SNV	het
PRKCH	protein kinase C, eta	uc010tsb.1:c.A505C:p.T169P	exonic	nonsynonymous SNV	het

VTI1B	vesicle transport through interaction with t-SNAREs homolog 1B (yeast)	uc001xju.2:c.A147T:p.K49N	exonic	nonsynonymous SNV	het
ZFYVE26	zinc finger, FYVE domain containing 26	uc001xka.2:c.T2825C:p.M942T	exonic	nonsynonymous SNV	het
MYO5C	myosin VC	uc010bff.2:c.G95A:p.R32K	exonic	nonsynonymous SNV	het
ITGAX	integrin, alpha X (complement component 3 receptor 4 subunit)	uc002ebt.2:c.G3232A:p.V1078M	exonic	nonsynonymous SNV	het
FUK	fucokinase	uc002eyz.2:c.G521A:p.R174H	exonic	nonsynonymous SNV	het
MLKL	mixed lineage kinase domain-like	uc002fdb.2:c.C771A:p.F257L	exonic	nonsynonymous SNV	het
PKD1L2	polycystic kidney disease 1-like 2	uc002fgf.1:c.C259G:p.L87V	exonic	nonsynonymous SNV	het
TOP3A	topoisomerase (DNA) III alpha	uc010cpz.1:c.C649T:p.R217C	exonic	nonsynonymous SNV	het
ZBTB7A	zinc finger and BTB domain containing 7A	uc002lzh.2:c.616_624del:p.206_208del	exonic	nonframeshift deletion	het
MPHOSPH10	M-phase phosphoprotein 10 (U3 small nucleolar ribonucleoprotein)	uc002sht.1:c.2035_2039del:p.679_680del	exonic	frameshift deletion	het
SEMA4F	sema domain, immunoglobulin domain (Ig), transmembrane domain (TM) and short cytoplasmic domain, (semaphorin) 4F	uc010ysb.1:c.G358C:p.G120R	exonic; splicing	nonsynonymous SNV	het
IMMT	inner membrane protein, mitochondrial	uc002sqz.3:c.G319A:p.G107S	exonic	nonsynonymous SNV	het
RANBP2	RAN binding protein 2	uc002tem.3:c.C8062G:p.L2688V	exonic	nonsynonymous	het

				SNV	
HSPDE1A	NA	uc002uou.2:c.A1346G:p.D449G	exonic	nonsynonymous SNV	het
KIAA0772	NA	uc011aah.1:c.C1066G:p.R356G	exonic	nonsynonymous SNV	het
COL20A1	collagen, type XX, alpha 1	uc011aav.1:c.G2584A:p.D862N	exonic	nonsynonymous SNV	het
CHRNA4	cholinergic receptor, nicotinic, alpha 4	uc010gke.1:c.C62T:p.S21F	exonic	nonsynonymous SNV	het
PCNT	pericentrin	uc002zji.3:c.C233A:p.T78N	exonic	nonsynonymous SNV	het
PCNT	pericentrin	uc002zji.3:c.A3008C:p.K1003T	exonic	nonsynonymous SNV	het
TOM1	target of myb1 (chicken)	uc011amk.1:c.C133T:p.H45Y	exonic	nonsynonymous SNV	het
THUMPD3	THUMP domain containing 3	uc003brn.3:c.C1334T:p.T445I	exonic	nonsynonymous SNV	het
ISY1	ISY1 splicing factor homolog (S. cerevisiae)	uc003elo.1:c.G499T:p.D167Y	exonic	nonsynonymous SNV	het
PAM	peptidylglycine alpha-amidating monooxygenase	uc003knz.2:c.C484T:p.R162W	exonic	nonsynonymous SNV	het
NOD1	nucleotide-binding oligomerization domain containing 1	uc003tav.2:c.C1909T:p.R637C	exonic	nonsynonymous SNV	het
NOM1	nucleolar protein with MIF4G domain 1	uc003wmy.2:c.T1321G:p.F441V	exonic	nonsynonymous SNV	het
FOXE1	forkhead box E1 (thyroid transcription factor 2)	uc004axu.2:c.532_537del:p.178_179del	exonic	nonframeshift deletion	het
PHPT1	phosphohistidine phosphatase 1	uc011mei.1:c.T308C:p.L103P	exonic	nonsynonymous SNV	het
DCAF8L2	DDB1 and CUL4 associated factor 8-like 2	uc011mjy.1:c.426_434del:p.142_145del	exonic	nonframeshift deletion	het

Appendix E: Family II

Table E2: Rare variants in batch 1

Gene	Full Gene Name		Func	Exonic Func	Call_343
CLCNKA	chloride channel Ka	uc001axu.2:c.G218A:p.C73Y	exonic	nonsynonymous SNV	het
HNRNPR	heterogeneous nuclear ribonucleoprotein R	uc009vql.2:c.T413C:p.I138T	exonic	nonsynonymous SNV	het
KIAA0319L	KIAA0319-like	uc001byw.2:c.1144_1146del:p.382_382del	exonic	nonframeshift deletion	het
FAM176B	family with sequence similarity 176, member B	uc001cai.1:c.C341G:p.T114R	exonic	nonsynonymous SNV	het
KIAA0754	KIAA0754	uc009vvt.1:c.G2156A:p.G719D	exonic	nonsynonymous SNV	het
TAL1	T-cell acute lymphocytic leukemia 1	uc009vyq.2:c.A329G:p.Q110R	exonic	nonsynonymous SNV	het
GNAI3	guanine nucleotide binding protein (G protein), alpha inhibiting activity polypeptide 3	uc001dxz.2:c.T1053A:p.C351X	exonic	stopgain SNV	het
HDGF	hepatoma-derived growth factor	uc009wsg.2:c.624_630del:p.208_210del	exonic	frameshift deletion	het
APOBEC4	apolipoprotein B mRNA editing enzyme, catalytic polypeptide-like 4 (putative)	uc001gqn.2:c.T328C:p.Y110H	exonic	nonsynonymous SNV	het

ASPM	asp (abnormal spindle) homolog, microcephaly associated (Drosophila)	uc001gtu.2:c.A988G:p.N330D	exonic	nonsynonymous SNV	het
TMEM206	transmembrane protein 206	uc001hjc.3:c.G337A:p.A113T	exonic	nonsynonymous SNV	het
GPR158	G protein-coupled receptor 158	uc001isk.2:c.G1210A:p.V404I	exonic	nonsynonymous SNV	het
RBP3	retinol binding protein 3, interstitial	uc001jez.2:c.G1840T:p.D614Y	exonic	nonsynonymous SNV	het
ATE1	arginyltransferase 1	uc010qts.1:c.G448T:p.A150S	exonic	nonsynonymous SNV	het
CHST15	carbohydrate (N-acetylgalactosamine 4-sulfate 6-O) sulfotransferase 15	uc001lhl.2:c.A27G:p.I9M	exonic	nonsynonymous SNV	het
FOXI2	forkhead box I2	uc009yas.2:c.G425C:p.S142T	exonic	nonsynonymous SNV	het
KNDC1	kinase non-catalytic C-lobe domain (KIND) containing 1	uc001llz.1:c.T5183A:p.V1728D	exonic	nonsynonymous SNV	het
CYP2R1	cytochrome P450, family 2, subfamily R, polypeptide 1	uc001mls.1:c.T686C:p.M229T	exonic	nonsynonymous SNV	het
INSC	inscuteable homolog (Drosophila)	uc001mmc.2:c.G1247A:p.R416Q	exonic	nonsynonymous SNV	het
CASP1	caspase 1, apoptosis-related cysteine peptidase	uc010rvi.1:c.C140T:p.S47F	exonic	nonsynonymous SNV	het

	(interleukin 1, beta, convertase)				
IPO8	importin 8	uc001rje.1:c.C459G:p.C153W	exonic	nonsynonymous SNV	het
CEP290	centrosomal protein 290kDa	uc001taq.2:c.T1058G:p.L353R	exonic	nonsynonymous SNV	het
UHRF1BP1L	UHRF1 binding protein 1-like	uc001tgp.2:c.C1202T:p.T401I	exonic	nonsynonymous SNV	het
ATP8A2	ATPase, aminophospholipid transporter, class I, type 8A, member 2	uc001uql.1:c.G25T:p.D9Y	exonic	nonsynonymous SNV	het
SOHLH2	spermatogenesis and oogenesis specific basic helix-loop-helix 2	uc001uvj.2:c.C581T:p.S194L	exonic	nonsynonymous SNV	het
FREM2	FRAS1 related extracellular matrix protein 2	uc001uwv.2:c.G3811A:p.E1271K	exonic	nonsynonymous SNV	het
ABCC4	ATP-binding cassette, sub-family C (CFTR/MRP), member 4	uc010afk.2:c.G2711A:p.R904H	exonic	nonsynonymous SNV	het
STRN3	striatin, calmodulin binding protein 3	uc001wqu.2:c.G287A:p.R96Q	exonic	nonsynonymous SNV	het
FBXO33	F-box protein 33	uc001wvk.2:c.T521G:p.L174R	exonic	nonsynonymous SNV	het
ZFYVE26	zinc finger, FYVE domain containing 26	uc001xka.2:c.C4502A:p.T1501K	exonic	nonsynonymous SNV	het
ZFYVE26	zinc finger, FYVE domain	uc001xka.2:c.G4358C:p.C1453S	exonic	nonsynonymous SNV	het

	containing 26				
MLH3	mutL homolog 3 (E. coli)	uc001xrd.1:c.T2608A:p.S870T	exonic	nonsynonymous SNV	het
MLH3	mutL homolog 3 (E. coli)	uc001xrd.1:c.A2504T:p.D835V	exonic	nonsynonymous SNV	het
PLDN	pallidin homolog (mouse)	uc001zvq.2:c.C224T:p.T75I	exonic; splicing	nonsynonymous SNV	het
EID1	EP300 interacting inhibitor of differentiation 1	uc001zxc.1:c.A172G:p.M58V	exonic	nonsynonymous SNV	het
AP4E1	adaptor-related protein complex 4, epsilon 1 subunit	uc001zyx.1:c.C1036A:p.L346M	exonic	nonsynonymous SNV	het
GCOM1	GRINL1A complex locus	uc002aei.2:c.G1182T:p.K394N	exonic	nonsynonymous SNV	het
NARG2	NMDA receptor regulated 2	uc002ago.2:c.C1745T:p.T582I	exonic	nonsynonymous SNV	het
ADAMTS7	ADAM metallopeptidase with thrombospondin type 1 motif, 7	uc002bej.3:c.C3164T:p.P1055L	exonic	nonsynonymous SNV	het
CCL22	chemokine (C-C motif) ligand 22	uc002elh.2:c.G250T:p.V84L	exonic	nonsynonymous SNV	het
CDH16	cadherin 16, KSP-cadherin	uc002eqm.2:c.G679C:p.V227L	exonic	nonsynonymous SNV	het
ZNF18	zinc finger protein 18	uc002gnh.1:c.G952A:p.E318K	exonic	nonsynonymous SNV	het
ACOX1	acyl-CoA oxidase 1, palmitoyl	uc002jqe.2:c.C1418T:p.T473I	exonic	nonsynonymous SNV	het
CCDC40	coiled-coil domain containing 40	uc002jxm.3:c.C829T:p.R277C	exonic	nonsynonymous SNV	het
MAN2B1	mannosidase,	uc002mub.2:c.G1345A:p.V449I	exonic	nonsynonymous	het

	alpha, class 2B, member 1			SNV	
CC2D1A	coiled-coil and C2 domain containing 1A	uc002mxq.1:c.A532G:p.M178V	exonic	nonsynonymous SNV	het
IL27RA	interleukin 27 receptor, alpha	uc002mxx.2:c.C1205T:p.A402V	exonic	nonsynonymous SNV	het
BRD4	bromodomain containing 4	uc002nar.2:c.C119T:p.A40V	exonic	nonsynonymous SNV	het
ZNF431	zinc finger protein 431	uc002npp.2:c.A840G:p.I280M	exonic	nonsynonymous SNV	het
C19orf12	chromosome 19 open reading frame 12	uc002nsj.2:c.C101T:p.A34V	exonic	nonsynonymous SNV	het
TSHZ3	teashirt zinc finger homeobox 3	uc002nsy.3:c.G76A:p.V26M	exonic	nonsynonymous SNV	het
DMKN	dermokine	uc002nzm.3:c.839_840insGCAGCAGTGGCA:p.G280delinsGSSGS	exonic	nonframeshift insertion	het
PLEKHG2	pleckstrin homology domain containing, family G (with RhoGef domain) member 2	uc010xva.1:c.C3247T:p.R1083W	exonic	nonsynonymous SNV	het
PRX	periaxin	uc002onq.2:c.C1058T:p.P353L	exonic	nonsynonymous SNV	het
QPCTL	glutaminyl-peptide cyclotransferase-like	uc010ekn.2:c.C74T:p.P25L	exonic	nonsynonymous SNV	het
ZSCAN4	zinc finger and SCAN domain containing 4	uc002qpu.2:c.97_98insCTG:p.V33delinsAV	exonic	nonframeshift insertion	het
NCOA1	nuclear receptor	uc002rfi.2:c.C713T:p.S238L	exonic	nonsynonymous	het

	coactivator 1			SNV	
NCOA1	nuclear receptor coactivator 1	uc002rfi.2:c.T785C:p.M262T	exonic	nonsynonymous SNV	het
SOS1	son of sevenless homolog 1 (Drosophila)	uc002rrl.2:c.T294A:p.D98E	exonic	nonsynonymous SNV	het
PCYOX1	prenylcysteine oxidase 1	uc002sgn.3:c.A629G:p.E210G	exonic	nonsynonymous SNV	het
TANK	TRAF family member-associated NFKB activator	uc002ubr.1:c.G420T:p.K140N	exonic	nonsynonymous SNV	het
LRP2	low density lipoprotein receptor-related protein 2	uc002ues.2:c.T2198C:p.M733T	exonic	nonsynonymous SNV	het
TTN	titin	uc002unb.2:c.T13589G:p.I4530R	exonic	nonsynonymous SNV	het
DNER	delta/notch-like EGF repeat containing	uc002vpv.2:c.323_324insGCA:p.C108delinsCS	exonic	nonframeshift insertion	het
SP140	SP140 nuclear body protein	uc002vqn.2:c.C963A:p.S321R	exonic	nonsynonymous SNV	het
ATRN	attractin	uc002wil.2:c.T1237C:p.S413P	exonic	nonsynonymous SNV	het
MYH7B	myosin, heavy chain 7B, cardiac muscle, beta	uc002xbi.1:c.G3352A:p.A1118T	exonic	nonsynonymous SNV	het
SLC32A1	solute carrier family 32 (GABA vesicular transporter), member 1	uc002xjc.2:c.C838G:p.L280V	exonic	nonsynonymous SNV	het
PRODH	proline dehydrogenase	uc002zoj.3:c.G1135A:p.A379T	exonic	nonsynonymous SNV	het

	(oxidase) 1				
CHL1	cell adhesion molecule with homology to L1CAM (close homolog of L1)	uc011asi.1:c.G127T:p.V43F	exonic	nonsynonymous SNV	het
ITPR1	inositol 1,4,5-trisphosphate receptor, type 1	uc010hca.1:c.G3127A:p.G1043S	exonic	nonsynonymous SNV	het
DKFZp434B103	NA	uc003btk.2:c.G136A:p.E46K	exonic; splicing	nonsynonymous SNV	het
TPRXL	NA	uc003byg.2:c.G527A:p.G176D	exonic	nonsynonymous SNV	het
NAT6	N-acetyltransferase 6 (GCN5-related)	uc003czi.2:c.C331A:p.H111N	exonic	nonsynonymous SNV	het
FLNB	filamin B, beta	uc003djl.2:c.G3884C:p.G1295A	exonic; splicing	nonsynonymous SNV	het
DZIP3	DAZ interacting protein 3, zinc finger	uc011bhm.1:c.C797T:p.A266V	exonic	nonsynonymous SNV	het
BOC	Boc homolog (mouse)	uc003eaa.1:c.G118A:p.V40I	exonic	nonsynonymous SNV	het
UROC1	urocanase domain containing 1	uc003eiz.1:c.C1504T:p.R502W	exonic	nonsynonymous SNV	het
DKFZp451E113	NA	uc003eqy.1:c.G389T:p.G130V	exonic	nonsynonymous SNV	het
FAIM	Fas apoptotic inhibitory molecule	uc003esr.2:c.A352G:p.M118V	exonic	nonsynonymous SNV	het
XRN1	5'-3' exoribonuclease 1	uc010huu.2:c.G3068C:p.G1023A	exonic	nonsynonymous SNV	het
PIGZ	phosphatidylinositol glycan anchor biosynthesis, class	uc003fxh.2:c.T1178C:p.I393T	exonic	nonsynonymous SNV	het

	Z				
RGS12	regulator of G-protein signaling 12	uc003ggx.1:c.A188T:p.Q63L	exonic	nonsynonymous SNV	het
RGS12	regulator of G-protein signaling 12	uc003ggx.1:c.A345T:p.E115D	exonic	nonsynonymous SNV	het
RGS12	regulator of G-protein signaling 12	uc003ggx.1:c.G356T:p.W119L	exonic	nonsynonymous SNV	het
RGS12	regulator of G-protein signaling 12	uc003ggv.2:c.G4125C:p.R1375S	exonic	nonsynonymous SNV	het
MSX1	msh homeobox 1	uc003gif.2:c.750_751insACCGGCTGCC:p.A250fs	exonic	frameshift insertion	het
PROM1	prominin 1	uc003gos.2:c.C2067A:p.S689R	exonic	nonsynonymous SNV	het
LRRC66	leucine rich repeat containing 66	uc003gzi.2:c.1217_1218insA:p.Q406fs	exonic	frameshift insertion	het
KDR	kinase insert domain receptor (a type III receptor tyrosine kinase)	uc003has.2:c.C1844T:p.T615I	exonic	nonsynonymous SNV	het
ADAMTS3	ADAM metallopeptidase with thrombospondin type 1 motif, 3	uc003hgk.1:c.C331T:p.H111Y	exonic	nonsynonymous SNV	het
SDAD1	SDA1 domain containing 1	uc011cbr.1:c.G1760A:p.R587Q	exonic	nonsynonymous SNV	het
CCDC111	coiled-coil domain containing 111	uc003iwn.2:c.646_649del:p.216_217del	exonic	frameshift deletion	het
RAI14	retinoic acid induced 14	uc010iur.2:c.G2203A:p.E735K	exonic	nonsynonymous SNV	het

GFRA3	GDNF family receptor alpha 3	uc003lco.2:c.C683T:p.P228L	exonic	nonsynonymous SNV	het
PCDHA13	NA	uc003lid.2:c.C674T:p.T225I	exonic	nonsynonymous SNV	het
PCDHB10	NA	uc003lix.2:c.G1618T:p.A540S	exonic	nonsynonymous SNV	het
PCDHGB5	NA	uc003lkf.1:c.G1200C:p.L400F	exonic	nonsynonymous SNV	het
PDGFRB	platelet-derived growth factor receptor, beta polypeptide	uc003lro.2:c.C2243T:p.S748L	exonic	nonsynonymous SNV	het
RASGEF1C	RasGEF domain family, member 1C	uc003mlp.3:c.A305C:p.K102T	exonic	nonsynonymous SNV	het
HIST1H4G	histone cluster 1, H4g	uc003nhf.2:c.T156G:p.Y52X	exonic	stopgain SNV	het
ZNF323	zinc finger protein 323	uc003nla.2:c.G155C:p.G52A	exonic	nonsynonymous SNV	het
TRIM10	tripartite motif containing 10	uc003npn.2:c.G756A:p.T252T	exonic; splicing	synonymous SNV	het
GNL1	guanine nucleotide binding protein-like 1	uc011dmi.1:c.C366A:p.D122E	exonic	nonsynonymous SNV	het
C6orf127	chromosome 6 open reading frame 127	uc003old.3:c.T148C:p.C50R	exonic	nonsynonymous SNV	het
MTCH1	mitochondrial carrier 1	uc011dtt.1:c.A587G:p.Y196C	exonic	nonsynonymous SNV	het
FGD2	FYVE, RhoGEF and PH domain containing 2	uc003onf.2:c.A4C:p.K2Q	exonic	nonsynonymous SNV	het
COL21A1	collagen, type XXI, alpha 1	uc003pcs.2:c.G787T:p.V263F	exonic	nonsynonymous SNV	het

KIAA2023	NA	uc003qld.2:c.G867C:p.Q289H	exonic	nonsynonymous SNV	het
IYD	iodotyrosine deiodinase	uc003qnu.1:c.G745A:p.V249M	exonic	nonsynonymous SNV	het
UNC84A	NA	uc010ksa.1:c.T50A:p.F17Y	exonic	nonsynonymous SNV	het
DNAJC30	DnaJ (Hsp40) homolog, subfamily C, member 30	uc003tys.1:c.A652G:p.I218V	exonic	nonsynonymous SNV	het
ABHD11	abhydrolase domain containing 11	uc003tzb.2:c.C41A:p.T14N	exonic	nonsynonymous SNV	het
CLIP2	CAP-GLY domain containing linker protein 2	uc003uao.2:c.G284A:p.R95Q	exonic	nonsynonymous SNV	het
CCL24	chemokine (C-C motif) ligand 24	uc011kga.1:c.C293T:p.P98L	exonic	nonsynonymous SNV	het
SEMA3A	sema domain, immunoglobulin domain (Ig), short basic domain, secreted, (semaphorin) 3A	uc003uhz.2:c.G247A:p.V83I	exonic	nonsynonymous SNV	het
ZNHIT1	zinc finger, HIT-type containing 1	uc003uye.2:c.C415T:p.R139C	exonic	nonsynonymous SNV	het
CHCHD3	coiled-coil-helix-coiled-coil-helix domain containing 3	uc003vre.2:c.A355G:p.K119E	exonic	nonsynonymous SNV	het
TEX15	testis expressed 15	uc003xil.2:c.G2170A:p.A724T	exonic	nonsynonymous SNV	hom
LY6K	lymphocyte antigen 6	uc011ljv.1:c.G302A:p.R101Q	exonic	nonsynonymous SNV	het

	complex, locus K				
NFX1	nuclear transcription factor, X-box binding 1	uc003zsp.1:c.G2989T:p.V997F	exonic	nonsynonymous SNV	het
CRB2	crumbs homolog 2 (Drosophila)	uc004bnw.1:c.G1174A:p.D392N	exonic	nonsynonymous SNV	het
INPP5E	inositol polyphosphate-5-phosphatase, 72 kDa	uc004cho.2:c.C331T:p.P111S	exonic	nonsynonymous SNV	het
EGFL6	EGF-like-domain, multiple 6	uc004cvi.2:c.G1406A:p.R469Q	exonic	nonsynonymous SNV	het
ASB9	ankyrin repeat and SOCS box containing 9	NA	splicing	NA	het
CXorf15	NA	uc010ney.1:c.G709T:p.A237S	exonic	nonsynonymous SNV	het

Appendix E: Family III

Table E3: Rare variants in batch 1

Gene	Full Gene Name		Func	Exonic Func	Call
CTBS	chitobiase, di-N-acetyl-	uc001dka.2:c.67_75del:p.23_25del	exonic	nonframeshift deletion	hom
PLEKHH2	pleckstrin homology domain containing, family H (with MyTH4 domain) member 2	uc002rte.3:c.2334_2335insTT:p.F778fs	exonic	frameshift insertion	hom
MFI2	antigen p97 (melanoma associated) identified by monoclonal antibodies 133.2 and 96.5	uc003fxk.3:c.G1192A:p.V398M	exonic	nonsynonymous SNV	hom
HTT	huntingtin	uc011bvq.1:c.110_111insGCAGCAGCAGCA:p.Q37delinsQQQQQ	exonic	nonframeshift insertion	hom
ZSCAN16	zinc finger and SCAN domain containing 16	uc003nkm.2:c.C703T:p.R235X	exonic	stopgain SNV	hom
MAS1L	MAS1 oncogene-like	uc011dlq.1:c.C167T:p.T56M	exonic	nonsynonymous SNV	hom
CLIC1	chloride intracellular channel 1	uc003nwr.2:c.G467A:p.S156N	exonic	nonsynonymous SNV	hom
FGD2	FYVE, RhoGEF and PH domain	uc010jwp.1:c.G359A:p.R120H	exonic	nonsynonymous SNV	hom

	containing 2				
TULP4	tubby like protein 4	uc003qrf.2:c.A4552G:p.I1518V	exonic	nonsynonymous SNV	hom
SMOC2	SPARC related modular calcium binding 2	uc003qwr.1:c.T681A:p.C227X	exonic	stopgain SNV	hom
TBP	TATA box binding protein	uc011ehf.1:c.221_222insGCA:p.Q74delinsQQ	exonic	nonframeshift insertion	hom
INHBA	inhibin, beta A	uc003thq.2:c.G821A:p.G274E	exonic	nonsynonymous SNV	hom
BLVRA	biliverdin reductase A	uc003tir.2:c.G814A:p.A272T	exonic	nonsynonymous SNV	hom
REEP4	receptor accessory protein 4	uc003xau.1:c.C661T:p.R221C	exonic	nonsynonymous SNV	hom
SMARCA2	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily a, member 2	uc003zhc.2:c.705_707del:p.235_236del	exonic	nonframeshift deletion	hom
NCOR2	nuclear receptor corepressor 2	uc010tba.1:c.1495_1496insCAGCAG:p.Q499delinsQQQ	exonic	nonframeshift insertion	hom
RNASE12	ribonuclease, RNase A family, 12 (non-active)	uc001vxt.2:c.T242A:p.I81N	exonic	nonsynonymous SNV	hom
SALL2	sal-like 2 (Drosophila)	uc001wbe.2:c.T2965C:p.S989P	exonic	nonsynonymous SNV	hom
CGRRF1	cell growth regulator with ring finger	uc001xay.2:c.A440G:p.K147R	exonic	nonsynonymous SNV	hom

	domain 1				
PAPLN	papilin, proteoglycan-like sulfated glycoprotein	uc001xnw.3:c.C1588A:p.P530T	exonic	nonsynonymous SNV	hom
VPS18	vacuolar protein sorting 18 homolog (S. cerevisiae)	uc001zne.2:c.G143A:p.R48H	exonic	nonsynonymous SNV	hom
MTFMT	mitochondrial methionyl-tRNA formyltransferase	uc002aof.3:c.T476C:p.V159A	exonic	nonsynonymous SNV	hom
C15orf39	chromosome 15 open reading frame 39	uc002azp.3:c.A1123G:p.T375A	exonic	nonsynonymous SNV	hom
KRTAP9-8	keratin associated protein 9-8	uc002hwh.3:c.C7T:p.H3Y	exonic	nonsynonymous SNV	hom
GPATCH8	G patch domain containing 8	uc002igv.1:c.G3940A:p.G1314S	exonic	nonsynonymous SNV	hom
LRRC46	leucine rich repeat containing 46	uc002ima.2:c.G322A:p.D108N	exonic	nonsynonymous SNV	hom
ZNF652	zinc finger protein 652	uc002iov.3:c.T56C:p.V19A	exonic	nonsynonymous SNV	hom
ANKFN1	ankyrin-repeat and fibronectin type III domain containing 1	uc002iun.1:c.C218T:p.T73M	exonic	nonsynonymous SNV	hom
REXO1	REX1, RNA exonuclease 1 homolog (S. cerevisiae)	uc010dsq.2:c.G607A:p.E203K	exonic	nonsynonymous SNV	hom
ZFR2	zinc finger RNA	uc002lyw.2:c.C2303T:p.S768F	exonic	nonsynonymous	hom

	binding protein 2			SNV	
LHB	luteinizing hormone beta polypeptide	uc002plt.2:c.G241T:p.D81Y	exonic	nonsynonymous SNV	hom
KLK10	kallikrein-related peptidase 10	uc002puy.2:c.818_912del:p.273_277del	exonic	frameshift deletion	hom
NLRP7	NLR family, pyrin domain containing 7	uc002qig.3:c.C2579G:p.T860S	exonic	nonsynonymous SNV	hom
C20orf26	chromosome 20 open reading frame 26	uc002wru.2:c.G1657A:p.E553K	exonic	nonsynonymous SNV	hom
DNMT3B	DNA (cytosine-5-)-methyltransferase 3 beta	uc002wyg.2:c.C758T:p.T253M	exonic	nonsynonymous SNV	hom
NCOA6	nuclear receptor coactivator 6	uc002xaw.2:c.G3262T:p.V1088L	exonic	nonsynonymous SNV	hom
MAFB	v-maf musculoaponeurotic fibrosarcoma oncogene homolog B (avian)	uc002xji.2:c.A542T:p.Q181L	exonic	nonsynonymous SNV	hom
PRKWNK4	NA	uc010wgy.1:c.G343A:p.G115R	exonic	nonsynonymous SNV	hom
RPL13AP6	NA	NA	splicing	NA	hom
KIAA0737	NA	uc010tlv.1:c.C1265T:p.S422F	exonic	nonsynonymous SNV	hom
LOC100130015	NA	uc010cjd.2:c.C503G:p.T168S	exonic	nonsynonymous SNV	hom
TSPYL3	NA	uc002wxm.1:c.C266T:p.T89I	exonic	nonsynonymous SNV	hom
PTPRV	NA	uc009xaa.2:c.385_386insCCTCGCTCCTCGCTCCTCGCTCCTCGCT:p.X129delinsSSLLAPRSSL	exonic	stoploss SNV	hom
DKFZp686C09209	NA	uc010vpv.1:c.G210C:p.K70N	exonic	nonsynonymous SNV	hom

Appendix F: Table F1: Intra -reliability test

	Mean	Std. Deviation	means dif.	T-test P-value	R correlation coeff.	ICCC
T17_M_D_1	9.677	0.448	-0.009	0.746	0.990	0.981
T17_M_D_2	9.686	0.469				
T17_B_L_1	10.754	0.639	0.010	0.376	0.999	0.998
T17_B_L_2	10.744	0.641				
T16_M_D_1	10.071	0.601	-0.013	0.404	1.000	0.999
T16_M_D_2	10.084	0.604				
T16_B_L_1	11.166	0.499	-0.005	0.378	1.000	0.999
T16_B_L_2	11.171	0.506				
T15_M_D_1	6.524	0.362	0.010	0.738	0.999	0.998
T15_M_D_2	6.514	0.356				
T15_B_L_1	9.171	0.447	0.082	0.350	0.809	0.654
T15_B_L_2	9.089	0.374				
T14_M_D_1	6.647	0.358	-0.013	0.388	0.999	0.997
T14_M_D_2	6.660	0.351				
T14_B_L_1	8.855	0.459	0.002	0.723	0.999	0.999
T14_B_L_2	8.853	0.462				
T13_M_D_1	7.582	0.336	-0.004	0.459	0.999	0.997
T13_M_D_2	7.586	0.343				
T13_B_L_1	7.363	1.073	0.003	0.766	1.000	0.999
T13_B_L_2	7.361	1.067				
T12_M_D_1	6.467	0.367	0.002	0.818	0.998	0.996
T12_M_D_2	6.465	0.372				
T12_B_L_1	5.447	0.619	0.005	0.429	0.999	0.999
T12_B_L_2	5.442	0.625				
T11_M_D_1	8.446	0.418	-0.005	0.438	0.999	0.998
T11_M_D_2	8.451	0.411				

Table F1: Intra -reliability test (continued)

	Mean	Std. Deviation	means dif.	T-test P-value	R correlation coeff.	ICCC
T11_B_L_1	6.196	0.591	0.001	0.914	0.999	0.998
T11_B_L_2	6.195	0.597				
T21_M_D_1	8.514	0.395	0.002	0.718	0.999	0.997
T21_M_D_2	8.512	0.406				
T21_B_L_1	6.272	0.526	-0.005	0.575	0.998	0.997
T21_B_L_2	6.277	0.531				
T22_M_D_1	6.538	0.470	0.001	0.941	0.997	0.994
T22_M_D_2	6.538	0.477				
T22_B_L_1	5.353	0.654	-0.019	0.578	0.985	0.971
T22_B_L_2	5.373	0.676				
T23_M_D_1	7.474	0.497	-0.004	0.546	0.999	0.998
T23_M_D_2	7.478	0.502				
T23_B_L_1	7.079	0.860	0.067	0.281	0.972	0.944
T23_B_L_2	7.013	0.843				
T24_M_D_1	6.778	0.352	0.003	0.633	0.998	0.996
T24_M_D_2	6.775	0.350				
T24_B_L_1	8.835	0.400	0.002	0.910	0.994	0.988
T24_B_L_2	8.833	0.420				
T25_M_D_1	6.480	0.317	-0.007	0.507	0.995	0.990
T25_M_D_2	6.487	0.327				
T25_B_L_1	8.985	0.311	-0.007	0.517	0.994	0.988
T25_B_L_2	8.992	0.307				
T26_M_D_1	10.095	0.584	0.083	0.286	0.939	0.881
T26_M_D_2	10.012	0.684				
T26_B_L_1	11.145	0.591	-0.070	0.197	0.959	0.919
T26_B_L_2	11.215	0.578				

Table F1: Intra -reliability test (continued)

	Mean	Std. Deviation	means dif.	T-test P-value	R correlation coeff.	ICCC
T27_M_D_1	9.505	0.436	0.000	1.000	0.992	0.983
T27_M_D_2	9.505	0.447				
T27_B_L_1	10.863	0.282	-0.018	0.588	0.953	0.908
T27_B_L_2	10.880	0.285				
T31_M_D_1	5.229	0.334	0.008	0.577	0.991	0.982
T31_M_D_2	5.222	0.330				
T31_B_L_1	5.550	0.267	0.000	1.000	0.985	0.971
T31_B_L_2	5.550	0.266				
T32_M_D_1	5.915	0.350	0.009	0.046	0.994	0.988
T32_M_D_2	5.906	0.358				
T32_B_L_1	5.829	0.462	-0.023	0.585	0.951	0.905
T32_B_L_2	5.853	0.457				
T33_M_D_1	6.659	0.474	0.023	0.488	0.973	0.948
T33_M_D_2	6.637	0.457				
T33_B_L_1	6.804	0.650	0.028	0.156	0.995	0.991
T33_B_L_2	6.776	0.632				
T34_M_D_1	6.878	0.375	0.011	0.274	0.996	0.992
T34_M_D_2	6.867	0.375				
T34_B_L_1	7.408	0.280	-0.010	0.425	0.991	0.982
T34_B_L_2	7.418	0.260				
T35_M_D_1	7.061	0.204	0.019	0.201	0.982	0.965
T35_M_D_2	7.042	0.221				
T35_B_L_1	8.067	0.407	0.018	0.156	0.996	0.992
T35_B_L_2	8.049	0.400				
T36_M_D_1	10.648	0.603	0.011	0.526	0.996	0.992
T36_M_D_2	10.637	0.589				
T36_B_L_1	10.167	0.471	-0.045	0.147	0.979	0.958

Table F1: Intra -reliability test (continued)

	Mean	Std. Deviation	means dif.	T-test P-value	R correlation coeff.	ICCC
T36_B_L_2	10.213	0.459				
T37_M_D_1	9.618	0.541	0.033	0.398	0.987	0.974
T37_M_D_2	9.585	0.523				
T37_B_L_1	10.037	0.670	-0.032	0.209	0.997	0.994
T37_B_L_2	10.068	0.660				
T41_M_D_1	5.270	0.244	0.017	0.490	0.947	0.896
T41_M_D_2	5.287	0.245				
T41_B_L_1	5.565	0.553	0.024	0.065	0.997	0.995
T41_B_L_2	5.541	0.559				
T42_M_D_1	5.813	0.352	-0.007	0.745	0.984	0.968
T42_M_D_2	5.819	0.375				
T42_B_L_1	5.888	0.500	0.010	0.634	0.990	0.981
T42_B_L_2	5.878	0.483				
T43_M_D_1	6.548	0.445	0.007	0.653	0.994	0.987
T43_M_D_2	6.542	0.441				
T43_B_L_1	6.924	0.462	0.011	0.655	0.988	0.976
T43_B_L_2	6.913	0.497				
T44_M_D_1	6.868	0.348	0.000	1.000	0.982	0.964
T44_M_D_2	6.868	0.365				
T44_B_L_1	7.546	0.369	0.009	0.637	0.986	0.972
T44_B_L_2	7.537	0.343				
T45_M_D_1	7.199	0.395	-0.016	0.549	0.989	0.978
T45_M_D_2	7.214	0.344				
T45_B_L_1	8.304	0.196	-0.018	0.202	0.981	0.962
T45_B_L_2	8.322	0.190				
T46_M_D_1	10.660	0.475	-0.051	0.082	0.983	0.966
T46_M_D_2	10.711	0.471				
T46_B_L_1	10.311	0.510	0.039	0.178	0.986	0.972
T46_B_L_2	10.272	0.530				
T47_M_D_1	9.856	0.413	-0.037	0.102	0.993	0.987
T47_M_D_2	9.893	0.393				
T47_B_L_1	9.401	0.691	-0.019	0.174	0.999	0.999
T47_B_L_2	9.420	0.672				